Università degli studi di Roma "La Sapienza"

DOTTORATO DI RICERCA IN

BIOLOGIA CELLULARE E DELLO SVILUPPO

Ciclo XXV

Identification and functional characterization of surface and secreted proteins of *Clostridium difficile*

Presentata da: Valeria Cafardi

Coordinatore Dottorato

Prof. Rodolfo Negri

Relatori

Dott. Davide Serruto

Prof. ssa Bianca Colonna

Esame finale anno 2013

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

a.a. - amino acid

ADP - adenosine diphosphate

BHI - brain heart infusion

DTT - dithiothreitol

EDTA - ethylenediaminetetraacetic acid

FRET - Fluorescence Resonance Energy Transfer

KDa - kilodalton

MW - molecular weight

OD - optical density

PAGE - polyacrylamide gel electrophoresis

PBS - phosphate buffered saline

SDS - sodium dodecyl sulphate

RT – room temperature

TCA - trichloroacetic acid

ABSTRACT

Clostridium difficile is a major cause of infectious diarrhea in hospitals worldwide. Recurrent infections and an increase in antibiotic-resistant strains have made treatment of *C. difficile* infections extremely difficult. Although the clostridial cell surface is recognized to play an important role in clostridial pathogenesis, biological functions of few cell surface proteins are known. Also, apart from the toxins, proteins released by *C. difficile* are poorly studied.

In order to comprehensively identify proteins secreted by *C. difficile* (either exposed on the surface or released in the extracellular environment) we tried two different proteomic approaches. Characterization of the bacterial surface can be done applying a 'surfome' analysis, which involves surface proteolysis of intact cells and identification of released peptides by mass spectrometry; the released proteins can instead be identified by performing similar analysis on bacterial culture supernatants. While the surface digestion approach was not successful because of bacterial lysis during sample preparation, analysis of culture media led to the identification of several proteins. We analyzed bacterial culture supernatants prepared from two clinical isolates, 630 and R20291. Interestingly, majority of the proteins identified were putative surface-associated proteins. These could be largely classified into proteins associated to the peptydoglycan (CWPs and LPXTG-anchored), extracellular hydrolases, transporters and flagellar proteins. Several proteins of unknown function were also identified. Among them, we selected for further characterization two proteins potentially involved in *C. difficile* pathogenesis.

CD630_28300, a protein of unknown function, shares sequence similarity with the anthrax lethal factor, a known zinc-metallopeptidase. We demonstrated that the recombinant protein binds divalent cations and is able to cleave fibronectin in a zinc-dependent manner *in vitro* and on human fibroblasts. Therefore we named

this novel zinc-metalloprotease Zmp1. Using mutants of Zmp1, we identified residues important in zinc binding and enzymatic activity.

CD630_36690, a protein with similarity to a sporulation and germination factor of *Bacillus subtilis*, was also identified in the proteomic screen. It is a small protein expressed on the surface of vegetative cells of *C. difficile* strain 630. An isogenic mutant lacking CD630_36690 demonstrated a phenotype suggesting a role in sporulation or germination.

Employing a combination of proteomic and molecular approaches we have discovered novel extracellular factors that may be important in key steps of clostridial pathogenesis.

INTRODUCTION

1. *Clostridium difficile*

Clostridium difficile is a Gram-positive, spore-forming obligate anaerobe, belonging to the phylum Firmicutes. It is an intestinal pathogen of humans and animals and is recognized as the most common cause of nosocomial diarrhea (Poxton *et al.*, 2001; Rupnik *et al.*, 2009; Kuipers and Surawicz, 2008). *C. difficile* was first described in 1935 as part of the intestinal microflora in neonates. Although the severe form of *C. difficile* disease, pseudomembranous colitis, was first reported in 1893, it was not identified as the causative agent of human disease until 1978 (Bartlett, 1994). As for other Firmicutes (Bacilli and Clostridia), *C. difficile* can exist in two forms: the vegetative cell and the spore (Fritze, 2004) (Fig. 1). The vegetative cell is the metabolically active and toxigenic form, while the spore, metabolically dormant, is able to survive for long periods and is resistant to harsh physical or chemical treatments (Setlow, 2007). The spore is the transmissible form of the disease, contributes to survival of the organism in the host and is responsible for recurrence of disease (Lawley *et al.*, 2009; Vedantam *et al.*, 2012).

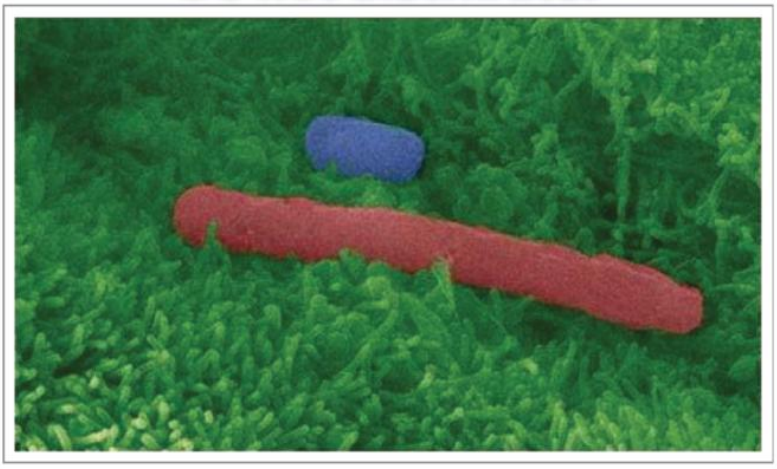


Figure 1. *C. difficile* spore and vegetative cell. From Vedantam *et al.*, 2012.

2. *C. difficile* associated disease (CDAD)

Upon ingestion of spores that are present in the air or on contaminated material, *C. difficile* can reach the gut. Presence in the gut can be asymptomatic (1–3% of healthy adults and 40–60% of neonates) or can produce a range of symptoms generally known as *Clostridium difficile* associated disease (CDAD). These symptoms usually appear in the presence of risk factors, such as prolonged antibiotic therapy, advanced age or hospitalization (Poxton *et al.,* 2001; Rupnik *et al.,* 2009; Kuipers and Surawicz, 2008).

Mild cases may present with slight fever, loose stools and abdominal cramps. The disease can progress to severe colitis with typical adherent pseudomembranes, which can coalesce to obscure the mucosa. Severe colitis can result in toxic megacolon, colon perforation and progressive multiorgan failure. Mortality in patients with CDAD ranges from 3–30% (Poxton *et al.*, 2001).

C. difficile is recognized as the main cause of infectious diarrhea that develops in patients after hospitalization and antibiotic treatment (Rupnik *et al.,* 2009) (Fig. 2). The association between antimicrobial therapy and CDAD has been almost universal, as *C. difficile* can only colonize the gut if the normal intestinal microbiota is disturbed or absent. Sequencing of the genome of the *C. difficile* strain 630, in 2006, showed that a large proportion (11%) of the genome consists of mobile genetic elements, mainly in the form of conjugative transposons. These mobile elements are putatively responsible for the acquisition by *C. difficile* of an extensive array of genes involved in antimicrobial resistance (Sebaihia *et al.,* 2006). Antibiotic treatment is always the leading risk for CDAD in people and animals. Initial studies focused on clindamycin and cephalosporins (Aronsson *et al.,* 1985; Gerding, 2009). Then, fluoroquinolones emerged as major inducing agents and were implicated in several outbreaks that could be controlled only by restraining or prohibiting use of the entire class (Gerding, 2004; Kallen *et al.,* 2009). Resistance to the antibiotics that are currently used to treat CDAD (metronidazole and vancomycin) has thus far not posed a significant threat.

Advanced age is a risk, with most reports showing sharp increases in incidence in persons over 65 years and a direct correlation with age above that threshold (Aronsson *et al.*, 1985; Pépin *et al.*, 2004). Another important risk is contact with the healthcare system, which is heavily contaminated by *C. difficile* (Best *et al.*, 2010; Cohen *et al.*, 2010; Donskey, 2010; Freeman *et al.*, 2010). Also, there is increasing evidence that proton pump inhibitors promote CDAD (Leonard *et al.*, 2007; Linsky *et al.*, 2010).

Patients in the community are also at risk for CDAD, albeit at a considerably lower rate than those who are hospitalized. Community-associated CDAD without previous direct or indirect contact with a hospital environment remains rare compared with hospital-acquired CDAD. Nevertheless, it has been reported in populations that were previously thought to be at low risk, such as young individuals and pregnant women (Rupnik *et al.*, 2009). Possible community sources for CDAD include soil, water, pets, animals used for food, meats and vegetables (al Saif *et al.*, 1996).

In animals, *C. difficile* was mainly known as an important pathogen in horses, although it has been reported to infect numerous wild and domestic animals (Rupnik, 2007). As with infection of humans, disease in animals is associated with non-protective normal gut flora (Songer and Anderson, 2006; Keel and Songer, 2006; Rodriguez-Palacios *et al.*, 2006).

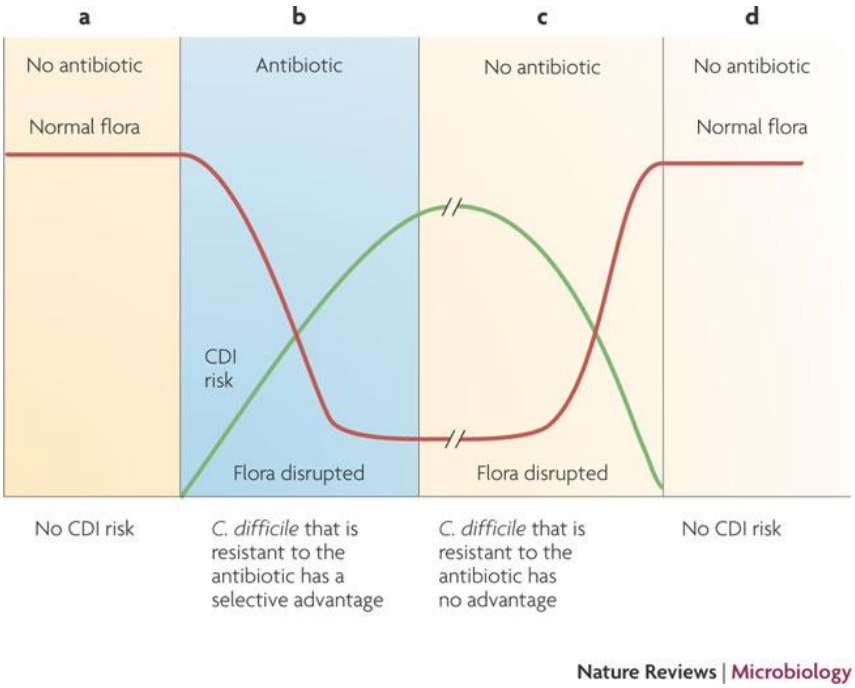


Figure 2. The effect of antibiotics on the normal gut flora and the risk of *Clostridium difficile* associated disease (CDAD). Patients are resistant to CDAD if their normal gut flora is not disrupted by antibiotics (**a**). Once antibiotic treatment starts, infection with a *C. difficile* strain that is resistant to the antibiotic is more likely while the antibiotic is being administered, owing to the presence of the antibiotic in the gut (**b**). When the antibiotic treatment stops, the levels of the antibiotic in the gut diminish rapidly, but the microflora remains disturbed for a variable period of time (indicated by the break in the graph), depending on the antibiotic given. During this time, patients can be infected with either resistant or susceptible *C. difficile* (**c**). Finally, after the microflora recovers, colonization resistance to *C. difficile* is restored (**d**). From Rupnik *et al.*, 2009.

2.1. Therapy

Two of the most difficult challenges for CDAD treatment are the management of multiple recurrences and the management of fulminant or severe complicated CDAD. Patients with multiple recurrences of CDAD typically respond to treatment with vancomycin or metronidazole, but then diarrhea symptoms resume within days to weeks after treatment is stopped. Between 20% and 50% of these recurrences are caused by new *C. difficile* organisms, indicating reinfection rather than a relapse of the original infection (Johnson *et al.*, 1989; O'Neill *et al.*, 1991). No highly effective means to treat these multiple recurrences have been devised, and most are treated with prolonged pulse dosing of vancomycin in the hope of keeping *C. difficile* from regrowing while the

normal flora recovers. The most effective treatment for these patients has been replenishment of the normal bacterial flora with a faecal transplant delivered either by nasogastric tube or by enema (Aas *et al.,* 2003). No highly effective treatment has been found for severe complicated CDAD, and in some cases surgical removal of the colon can be the only remaining life-saving measure (Lamontagne *et al.,* 2007). Newer agents for treatment of multiple relapsing and fulminant CDAD are needed, and several approaches are under investigation, including vaccines, toxin-binding agents and passive antibodies.

2.2. Emergence of hypervirulent strains

Since 2000, there has been a dramatic increase in rates and severity of CDAD noted in North America and much of Europe. The first reports were from Canada in 2003, showing a fourfold increase in CDAD rates, and included reviews showing fluoroquinolones as the dominant inducing agents and 10-fold-higher rates in persons over 65 years. Similar studies in the United States showed an increase in CDAD rate and mortality (Dubberke *et al.,* 2010; Redelings *et al.,* 2007; Ricciardi *et al.,* 2007). European data also showed increased CDAD rates and multiple hospital-associated outbreaks (Freeman *et al.*, 2010).

The increasing incidence and severity of CDAD in North America and much of Europe has been attributed to the emergence of hypervirulent ribotype 027 (McDonald *et al.,* 2005; Loo *et al.,* 2005; O'Connor *et al.,* 2009; Warny *et al.,* 2005). The factor thought to contribute to incidence is the strain's resistance to fluoroquinolones, and enhanced virulence may reflect increased toxin production *in vitro* (Warny *et al.,* 2005). In addition, emerging strains belonging to PCR-ribotype 017 and 078, that are also associated with severe disease, have been isolated in parts of Asia and Europe (Drudy *et al.,* 2007; Kim *et al.,* 2008). There is currently a widespread interest in understanding the underlying factors that have led to the emergence of these hypervirulent *C. difficile* strains.

3. *C. difficile* infection

Upon contact with contaminated surfaces, *C. difficile* spores and vegetative cells can be ingested. Spores, which are able to survive the acid environment of the stomach, can reach the small intestine. Here they germinate in response to specific germination-promoting factors, resulting in a vegetative morphotype (Fig. 3). In susceptible hosts, this vegetative form will proliferate in the colon, producing toxins (if the strain is toxigenic) and resulting in the characteristic pathology (Vedantam *et al.*, 2012).

During CDAD, *C. difficile* vegetative cells initiate the process of sporulation (Fig. 3). Spores are responsible for transmission of the disease. Patients with CDAD episodes enter a highly contagious super-shedder state, excreting high levels of infectious spores to the environment. These spores are resistant to most conventional disinfection methods and can survive for long periods (Lawley *et al.*, 2009). Also, spores are responsible for recurrencies of CDAD. A fraction of spores remains adhered to the intestinal surface. These spores, which are not affected by antibiotic therapies commonly used for CDAD, then germinate, outgrow and recolonize the host gastrointestinal tract, causing relapse of disease (Sarker and Paredes-Sabja, 2012).

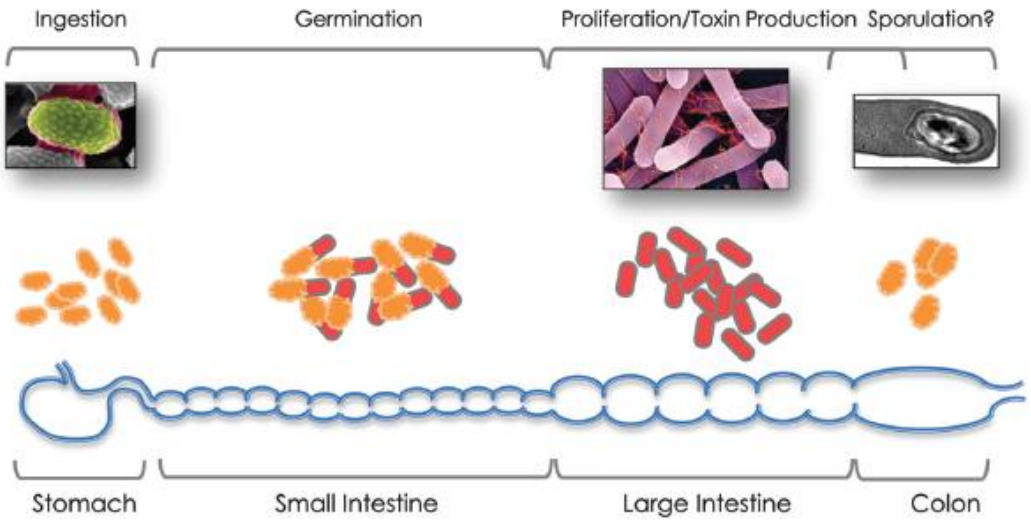


Figure 3. Schematic of human *C. difficile* infection. Modified from Vedantam *et al.*, 2012.

4. *C. difficile* virulence factors

Toxins are the most well studied virulence factors of *C. difficile*, as only toxigenic strains have been so far associated to disease (Kuehne *et al.*, 2010; Lyras *et al.*, 2009; Voth and Ballard, 2005; Schwan *et al.*, 2009). *C. difficile* produces two main toxins, toxin A and toxin B (Voth and Ballard, 2005), which account for symptoms of disease and produce intestinal damage, ensuring the release of nutrients from the damaged colonic epithelium. However, other factors must play a role in processes such as adaptation to the host's colonic environment, adhesion to the gut mucosa, colonization and dissemination of the infection.

Efforts to elucidate *C. difficile* pathogenic mechanisms have been hampered for many years by a lack of molecular tools (Kuehne *et al.*, 2011). In recent years, the sequencing of several *C. difficile* genomes (Sebaihia *et al.*, 2006; He *et al.*, 2013) and the spread of molecular tools that allow *C. difficile* genetic studies (Cartman *et al.*, 2012; Kuehne *et al.*, 2011) have led to a great wealth of knowledge regarding the molecular basis of *C. difficile* pathogenesis and nontoxin virulence factors likely involved in the early stages of CDAD. Still, very little is known about factors other than toxins involved in pathogenesis.

Bacterial adherence has been reported to play a role in *C. difficile* pathogenic strategy. It is clearly established that *C. difficile* can associate with intestinal mucosa in humans and hamsters (Borriello *et al.*, 1988). Also, there appears to be a positive correlation between virulence and mucosal adherence *in vivo* (Borriello *et al.*, 1988). Early studies of interactions between *C. difficile* vegetative cells and intestinal enterocyte cells demonstrated that *C. difficile* cells adhere to the microvilli of the apical surfaces of Caco-2 cells (Eveillard *et al.*, 1993; Cerquetti *et al.*, 2002). A recent study demonstrated that during infection of intestinal enterocyte-like Caco-2 cells, *C. difficile* differentially expressed approximately 7% of its genome, resulting in the upregulation of numerous putative auxiliary virulence factors (Janvilisri *et al.*, 2010).

Several surface factors involved in adhesion and colonization have been identified and partially characterized in the last years through *in vitro* or *in vivo* studies: most relevant are SLPs (Calabi *et al.*, 2002a), Cwp66 (Waligora *et al.*, 2001), Fbp68 (Barketi-Klai *et al.*, 2011), Cwp84 (Janoir *et al.*, 2007) and flagella (Dingle *et al.*, 2011; Tasteyre *et al.*, 2001a). Also, it is likely that during colonization of the gut *C. difficile* releases not only the known toxins but also other proteins. However, there is little information about proteins secreted by *C. difficile* into its environment. A recent report described Srl, a protein found in culture supernatants which modulates the cell sensitivity to toxins A and B (Miura *et al.*, 2011). Other factors which have been proposed to be involved in colonization are proteolytic enzymes (Poilane *et al.*, 2008) and capsule (Davies and Borriello, 1990), but these have been so far poorly characterized.

It is well known that formation of biofilms influences the ability of several pathogens to colonize and establish an infection (Nobbs *et al.*, 2009; Allsopp *et al.*, 2010), also providing an enclosed environment to escape immune responses and resist to antibiotics (Beloin *et al.*, 2008; Mah and O'Toole, 2001). Interestingly, *C. difficile* has been recently shown to form biofilms *in vitro*, an ability which could be crucial for clostridial colonization and persistence in the host (Dapa *et al.*, 2012; Dawson *et al.*, 2012).

4.1. Toxins

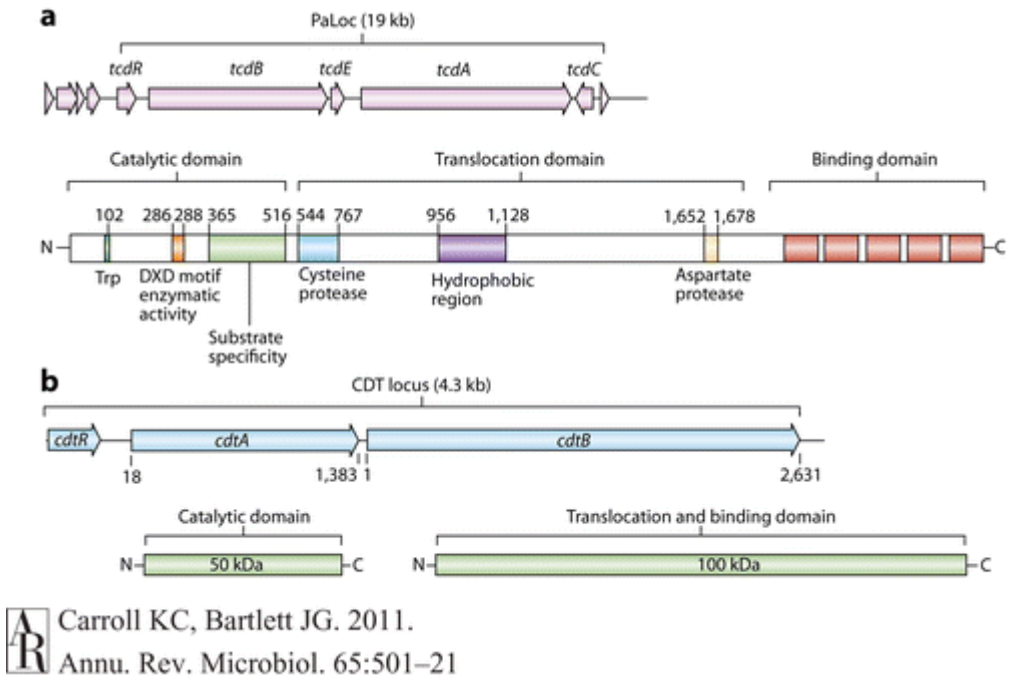
4.1.1. Toxin A and toxin B

The main virulence factors of *C. difficile* have been identified as toxins A and B, which target the Ras superfamily of GTPases for modification via glycosylation, ultimately leading to destruction of intestinal epithelial cells (Carter *et al.*, 2012; Voth and Ballard, 2005).

Toxin A and toxin B are encoded in the Pathogenicity Locus (PaLoc) (Fig. 4a). The PaLoc of *C. difficile* is approximately 19.6 kb in size and is stable and conserved in toxigenic strains (Cohen *et al.*, 2000; Voth and Ballard, 2005).

Nontoxigenic strains lack the PaLoc; however, isolates with a defective PaLoc can still cause disease (Cohen *et al.*, 2000; Curry *et al.*, 2007). Five genes are present on the PaLoc: *tcdA, tcdB, tcdC, tcdE* and *tcdR*. The two toxin genes, *tcdA* and *tcdB*, are closely aligned, separated by an intervening sequence (*tcdE*) (Voth and Ballard, 2005). *tcdE* encodes a holin, a protein whose pore-forming activity has been proposed to allow the release of TcdA and TcdB from the cell, although it has recently been shown that it does not exhibit pore-forming function in *C. difficile* (Olling *et al.*, 2012; Voth and Ballard, 2005). *tcdR*, found upstream of *tcdB*, is a major positive regulator of *tcdA* and *tcdB* expression, responsive to environmental conditions and increased during stationary phase (Voth and Ballard, 2005). *tcdC*, found downstream of *tcdA*, has been considered for long time a negative regulator of toxin production, but it has recently been shown that functionality of this gene does not affect toxin production (Cartman *et al.*, 2012; Carter *et al.*, 2012).

All the genes of the locus except *tcdC* are expressed during stationary phase. Dineen and co-workers have shown that toxin gene expression is also regulated by the global gene regulator CodY, which acts by monitoring environmental nutrient factors. The authors demonstrated that in the presence of sufficient nutrients CodY binds to the promoter region of *tcdR* and represses toxin gene expression. When nutrients in the environment are lacking, toxin gene expression is derepressed (Voth and Ballard, 2005; Carter *et al.*, 2012; Dineen *et al.*, 2007).



Carroll KC, Bartlett JG. 2011.
Annu. Rev. Microbiol. 65:501–21

Figure 4. Toxins produced by *C. difficile*. (**a**) Two large toxins, toxin A and toxin B (TcdA and TcdB), are encoded on the pathogenicity locus (PaLoc), which comprises five genes. Both toxins are single-chain proteins, and several functional domains and motifs have been identified. TcdB is shown in detail below the PaLoc. (**b**) A third toxin, the binary toxin or CDT (*C. difficile* transferase), is encoded on a separate region of the chromosome (CdtLoc) and comprises three genes. The binary toxin is composed of two unlinked proteins, CdtB and CdtA. CdtB has a binding function and CdtA is the enzymatic component. From Carroll and Bartlett, 2011.

TcdA and TcdB are part of the large clostridial toxin family along with *Clostridium sordellii* lethal toxin and hemorrhagic toxin and *Clostridium novyi* alpha toxin (Carter *et al.*, 2012). Both toxins are glucosyltransferases. They transfer a UDP-glucose to small GTPases, such as Rho, Rac, and Cdc42, in the cell. These small proteins are important in regulating signaling pathways. Glycosylation disrupts these pathways, which results in morphological changes, inhibition of cell division and membrane trafficking, and eventual cell death (Dawson *et al.*, 2009; Voth and Ballard, 2005).

TcdA binds to the apical side of the cell and, after internalization, causes cytoskeletal changes that result in disruption of tight junctions and loosening of the epithelial barrier, in cell death or in the production of inflammatory mediators that attract neutrophils. Disruption of tight junctions enables both TcdA and

TcdB to cross the epithelium. TcdB binds preferentially to the basolateral cell membrane. Both toxins are cytotoxic and induce the release of various immunomodulatory mediators from epithelial cells, phagocytes and mast cells, resulting in inflammation and the accumulation of neutrophils (Rupnik *et al.*, 2009) (Fig. 5).

TcdA (308 kDa), an enterotoxin, was originally believed to be the toxin associated with disease, and therefore necessary for virulence, until TcdA−, TcdB+ strains were associated with outbreaks of severe *C. difficile* infection (Freeman *et al.*, 2010). TcdB (270 kDa), a cytotoxin, is 100- to 1000-fold more toxic to culture cells than TcdA is. Lyras and co-workers reported that TcdB, not TcdA, is essential for virulence (Lyras *et al.*, 2009). However, a paper by Kuehne and co-workers refutes the assertion that only toxin B is essential for virulence and re-establishes the observation that both toxins can cause significant disease (Kuehne *et al.*, 2009). In that study, TcdA+, TcdB− isolates were as likely to cause disease as the wild-type strains.

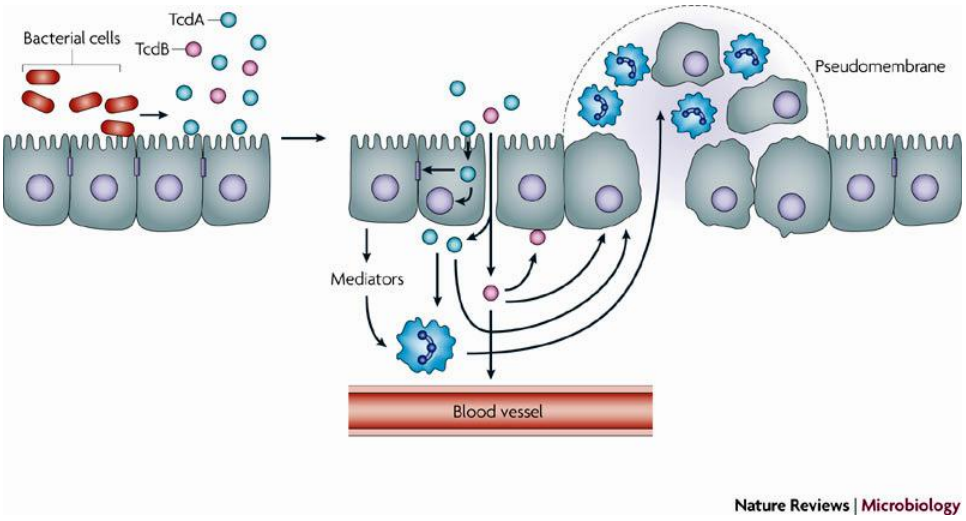


Figure 5. Clostridium difficile pathogenesis. *C. difficile* colonizes the intestine (colon) after disruption of the normal intestinal flora. In the schematic, bacterial cells are shown as free cells and attached to host cells. TcdA binds to the apical side of the cell and, after internalization, causes cytoskeletal changes that result in disruption of tight junctions and loosening of the epithelial barrier, in cell death or in the production of inflammatory mediators that attract neutrophils. Disruption of tight junctions enables both TcdA and TcdB to cross the epithelium. TcdB binds preferentially to the basolateral cell membrane. Both toxins are cytotoxic and induce the release of various immunomodulatory mediators from epithelial cells, phagocytes and mast cells, resulting in inflammation and the accumulation of neutrophils. From Rupnik *et al.*, 2009.

Both toxins are large single-stranded proteins, and recent X-ray crystallography and small angle X-ray scattering models (SAXS) of TcdB suggest four structural domains (Albesa-Jove *et al.,* 2010) (Fig. 4a). These domains include (*a*) a biologically active N-terminal glucosyltransferase protruding from the core of the protein; (*b*) a cysteine protease domain; (*c*) a middle translocation section that contains a hydrophobic region implicated in toxin delivery; and (*d*) a C-terminal receptor-binding domain (Albesa-Jove *et al.,* 2010). Toxin activity is located in the N-terminal domain. This portion is delivered into the cytosol of host cells (Jank and Aktories, 2008). Cleavage of the biologically active segment occurs by autoproteolysis via the cysteine protease domain (Albesa-Jove *et al.,* 2010). The C-terminal domain has short combined repetitive oligopeptides (CROPs) for receptor binding. In animal models of TcdA, carbohydrate structures play a role in toxin binding (Rupnik *et al.,* 2009; Voth and Ballard, 2005). These carbohydrates are not present in humans and the glycoprotein gp96 present in the human colon is the receptor for TcdA (Rupnik *et al.,* 2009).

Once inside the cell, the toxins target the Ras superfamily of small GTPases (Rho, Rac, and Cdc) (Dineen *et al.,* 2007), modifying them through glycosylation (Voth and Ballard, 2005). Glycosylation prevents the structural changes required for active conformation of these GTPases, significantly altering their function. Cells shrink and become rounded owing to disaggregation of the actin cytoskeleton, eventually dying (Schirmer and Aktories, 2004; Voth and Ballard, 2005). Tight junctions between epithelial cells are disrupted. This allows neutrophils to migrate to the intestines, contributing to the inflammatory response, typical of colitis (Voth and Ballard, 2005). Biological inactivation of the GTPases results in serious physiological consequences, such as inhibition of secretion, transcriptional regulation, and eventual apoptosis (Schirmer and Aktories, 2004; Voth and Ballard, 2005). In addition to direct cytotoxic effects, TcdA stimulates release of tumor necrosis factor from activated macrophages as

well as cytokine production. These activities cause fluid accumulation and further the inflammatory responses (Voth and Ballard, 2005).

4.1.2. *C. difficile* Transferase (Binary toxin CDT)

While the role of the toxins TcdA and TcdB in disease is well documented, the role of the binary toxin is less certain.

C. difficile transferase (CDT), also known as binary toxin, is encoded by the Cdt locus (CdtLoc) (Fig. 4b). It is found in approximately 6%–12.5% of strains overall (Schwan *et al.,* 2009; Sundriyal *et al.,* 2010). CDT is an ADP-ribosylating toxin that disrupts the cytoskeleton of the cell, leading to excessive fluid loss, rounding of the cell, and eventual cell death (Sundriyal *et al.,* 2010).

CDT is composed of two subunits, CDTa and CDTb (Fig. 4b). Each component alone is not cytotoxic, while together they cause cytotoxicity *in vitro*. The fact that the incidence of CDT is higher in some of the epidemic strains suggests that it contributes to the severity of disease (Sundriyal *et al.,* 2010; Geric *et al.,* 2006). Findings by Schwan and co-workers (Schwan *et al.,* 2009) revealed that CDT induces the formation of novel thin, dynamic, microtubules on the surface of epithelial cells, leading to increased adherence of bacteria *in vitro* and *in vivo*. Electron microscopy showed that these protrusions increase adherence to the epithelial cell surface by approximately fivefold *in vitro* and fourfold in the mouse large intestine, thereby possibly playing an important role in intestinal colonization (Schwan *et al.,* 2009).

4.2. Surface virulence factors

S-layer proteins

Many prokaryotes have a surface-associated crystalline or paracrystalline array, formed by identical subunits of glycoproteins or proteins, typically named surface-layer (S-layer) (Sara and Sleytr, 2000). Several functions have been proposed for bacterial S-layers, including acting as molecular sieves, protective

factors against parasitic attack, virulence factors and adhesion sites for extracellular proteins (Sara and Sleytr, 2000). *C. difficile* has two superimposed and structurally different S-layer lattices, a high-molecular-weight (HMW) SLP (~40 kDa) and a low-molecular-weight (LMW) SLP (~35 kDa) (Cerquetti *et al.*, 2000; Calabi *et al.*, 2001; Karjalainen *et al.*, 2001) (Fig. 6a).

A single gene, *slpA*, encodes both proteins produced through post-translational cleavage of a common precursor, with the N-terminal portion coding for the LMW SLP and the C-terminal portion coding for the HMW SLP (Calabi *et al.*, 2001). The translated gene product undergoes two rounds of post-translational cleavage, firstly to remove the signal sequence following secretion, and then internally to release the two mature SLPs (Fig. 6b).

Both HMW and LMW are tightly linked by a noncovalent complex (Fagan *et al.*, 2009). Interestingly, while the HMW SLP is localized in the internal surface of the bacteria and is immunologically conserved, the LMW SLP localizes to the external surface of the bacterium and is immunologically variable (Cerquetti *et al.*, 2000; Takeoka *et al.*, 1991; Fagan *et al.*, 2009), suggesting some role of LMW in evasion of the immune system (Vedantam *et al.*, 2012).

SLPs play a role in adhesion. Initial studies on the SLPs' adhesive properties demonstrated that removal of the SLPs through chemical treatments or treatment of *C. difficile* bacterial cells with anti-SLP antibodies abolished adherence of *C. difficile* to mouse 929 and human HeLa cells (Takeoka *et al.*, 1991). Later studies demonstrated that purified SLPs bind to intestinal tissues and several proteins of the extracellular matrix (Calabi *et al.*, 2002a), yet to date the precise host receptor has not been identified.

Besides the aforementioned *slpA*, *C. difficile* 630 genome contains 28 paralogs of the HMW SLP, which are known as the family of cell wall proteins (CWPs) (Sebaihia *et al.*, 2006). All of the members of the *C. difficile* CWP family contain up to three copies of the cell wall-binding motif (PF01422), a motif characteristic of the SLPs which is involved in anchoring the protein to the outer surface of the

bacterium, as well as a variable motif that might be function-specific (Sebaihia *et al.,* 2006). All the paralogs contain a signal sequence and are predicted to be surface localized; indeed, several have been detected experimentally (Wright *et al.,* 2005). Among these, some have been partially characterized and for them a role has been proposed in host cell adherence or other pathogenic processes.

Of the 29 paralogs identified so far, 12 map in a densely arranged cluster surrounding *slpA* and are all transcribed in the same direction, suggesting the possibility of coordinated regulation and related functions (Calabi *et al,* 2002a Karjalainen *et al,* 2001; Calabi *et al,* 2002b).

Despite the wealth of knowledge on a few CWPs, the lack of information on the remaining CWP paralogs hinders our ability to fully understand the role of these cell surface proteins in the pathogenesis of *C. difficile* and their importance for early colonization or evasion of the host's immune system. Indeed, evidence from microarray suggests that *C. difficile* remodels its cell surface during early contact with Caco-2 cells (Janvilisri *et al.,* 2010). These changes include the upregulation of several genes that encode putative auxiliary virulence factors, indicating that *C. difficile* remodels its envelope perhaps to enhance exposure of virulence attributes relevant for colonization, spread and/or early invasion stages (Janvilisri *et al.,* 2010).

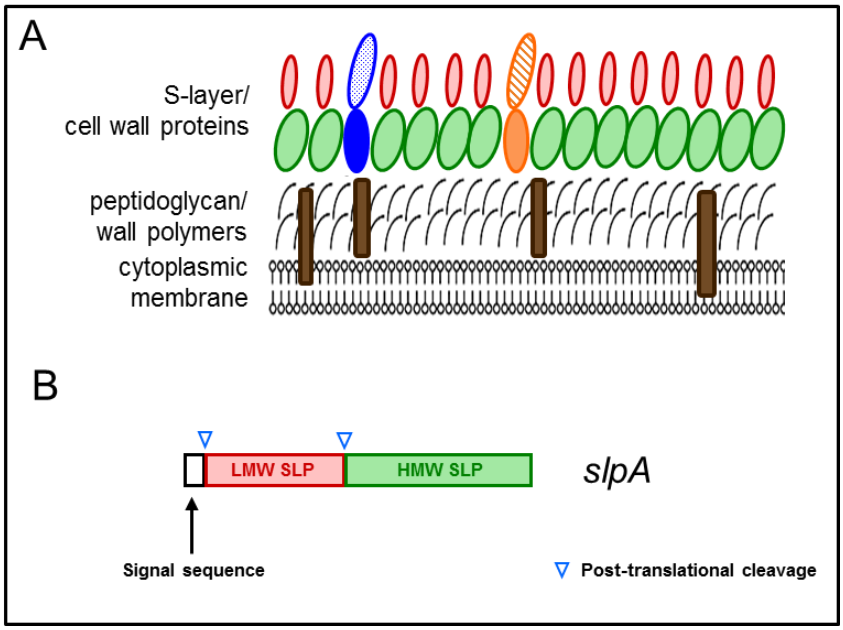


Figure 6. Model of the cell wall of *Clostridium difficile*. (**A**) The two SLPs are shown above the peptidoglycan layer: the HMW SLP (green) and the LMW SLP (red). Other minor cell wall proteins are shown as two-lobed structures (blue and orange). Putative cell wall polymers, including putative lipid-containing polymers, are shown as vertical brown bars. (**B**) The precursor protein SlpA, showing the cleavage sites generating the signal peptide and the mature HMW SLP and LMW SLP. Modified from Fagan *et al.*, 2009.

Cwp84 and Cwp13

During early colonization of host surfaces, in several bacterial species (i.e., *Streptococcus pyogenes* and *Pseudomonas aeruginosa*) proteolytic enzymes play an important role in degrading host proteins, contributing to nutrient acquisition, or in the processing of bacterial proteins involved in virulence (Maeda, 1996; Matsumoto, 2004).

The Cwp84 protease, conserved in various *C. difficile* strains (Savariau-Lacomme *et al.*, 2003) and localized in the locus containing the *slpA* gene, was detected by CDAD patient's antisera, suggesting a potential role of Cwp84 in pathogenesis of CDAD (Péchiné *et al.*, 2005a). Cwp84 is a cysteine protease, likely synthesized as an inactive proprotein of 80 kDa, that autoprocesses itself to a 47-kDa mature Cwp84 (Janoir *et al.*, 2007; Chapetonmontes *et al.*, 2011). Using both chemical and genetic techniques, the cysteine protease Cwp84 was shown to mediate cleavage of the mature SlpA precursor (Kirby *et al.*, 2009;

Dang *et al.,* 2010). In a *C. difficile cwp84* mutant strain, SlpA weakly binds to the bacteria's cell wall primarily due to the lack of formation of the mature subunits, which might explain the slower growth of a *C. difficile cwp84* strain when compared with parental strain (Kirby *et al.,* 2009). Neither chemical inhibition of Cwp84 (Dang *et al.,* 2010) nor inactivation of the cwp84 gene (Kirby *et al.,* 2009) resulted in lethality, though severe growth defects were seen in both cases. The presence of uncleaved SlpA in the cell wall of the cwp84 mutant resulted in aberrant retention of other cell wall proteins, suche as Cwp66 and Cwp2, at the cell surface and subsequent release in the culture supernatant (de la Riva *et al.,* 2011). These results indicate that correct processing of SlpA is important to retain healthy bacterial cells and suggest that perturbation of processing may affect the ability of bacteria to compete with other bacterial species in certain environments, for example, in the complex microbiota of the intestine. Despite these observations, *C. difficile cwp84* mutants are competent in causing CDAD (Kirby *et al.,* 2009).

Cwp84 is also involved in degradation of fibronectin, vironectin and laminin presumably inducing a loss in the integrity of the host's colonic epithelium and an increase in toxin diffusion (Janoir *et al.,* 2007). In recent years, Cwp84 has been considered as a potential target for chemical inhibitors (Dang *et al.,* 2010; Tam Dang *et al.,* 2012) and as a potential vaccine candidate (Sandolo *et al.,* 2011; Péchiné *et al.,* 2011).

Another putative cysteine protease, Cwp13, shares 63.2% amino acid identity to Cwp84, but plays an auxiliary role in the assembly of the S-layer (de la Riva *et al.,* 2011). Cwp13 was shown to cleave SlpA in a different site respect to Cwp84, an activity that was suggested to reflect a role in cleavage and degradation of misfolded proteins at the cell surface. It was also shown to have a role in processing Cwp84, although it is not essential for Cwp84 activity (de la Riva *et al.,* 2011).

Cwp66

Cwp66 has a molecular size of 66 kDa and is anchored to the cell surface through its N-terminal domain (Waligora *et al.,* 2001). Only the C-terminal of Cwp66 was detectable by immunoelectron microscopy after cells were subjected to a heat shock, suggesting that the N-terminal domain is embedded in the cell wall and is inaccessible to antibodies. Competitive inhibition assays showed that Cwp66 is an adhesin (Waligora *et al.,* 2001). However, given the fact that the C-terminal domain could not be detected by immunoelectorn microscopy of non-heat shock cells, the precise role of Cwp66 during *in vivo* adhesion in absence of heat shock remains unclear.

CwpV

CwpV has three cell wall anchoring motifs, a region of unknown function, and a C-terminal domain of variable number repeats (Emerson *et al.,* 2009). CwpV is regulated by phase variation through DNA inversion by a site-specific recombinase and is expressed in approximately 5% of the cells under laboratory conditions (Emerson *et al.,* 2009). Those cells expressing CwpV had no increment in adhesion to Caco-2 cells, but had the ability to autoaggregate forming denser and more randomly packed cellular organizations (Reynolds et *al.,* 2011). Although repeats vary in size and are antigenically different between strains, CwpV exhibited conserved aggregation promoting function organizations (Reynolds *et al.,* 2011). Autoaggregative proteins play a role in biofilm formation. Recently, *C. difficile* has been shown to form biofilms on abiotic substrates (Dapa *et al.,* 2012; Dawson *et al.,* 2012). Moreover, infected mice have been shown to have large aggregates of *C. difficile* cells in regions with severe inflammation (Lawley *et al.,* 2009). These data suggest that CwpV might enhance *C. difficile*'s ability to colonize colonic surfaces (Reynolds *et al.,* 2011). Due to its presence in a small fraction of the population, it is likely that CwpV might play a role in antigenic variation, or alternatively, might allow evasion of

the innate immune system of a small fraction to cause further relapse episodes (Reynolds *et al.,* 2011).

GroEL

Adherence to tissue culture cells was demonstrated to be augmented by various stresses, such as heat, osmotic and acid shock as well as iron insufficiency (Waligora *et al.,* 1999).

GroEL localizes at the cell surface, is predominantly membrane bound, and is released extracellularly after heat shock (Hennequin *et al.,* 2001b). Treatment of *C. difficile* vegetative cells with polyclonal antibodies specific to *C. difficile* GroEL reduced their ability to adhere to Vero cells in culture (Hennequin *et al.,* 2001b). The *C. difficile* GroEL is upregulated during heat shock (Hennequin *et al.,* 2001a), suggesting that it might play a role during clinically relevant heat stress as recently suggested (Jain *et al.,* 2011). However, genetic studies will be required to determine the precise role of GroEL in CDAD.

Flagella

Flagella have been implicated in internalization of *Campylobacter jejuni* and *Legionella pneumophila* (Dietrich *et al.,* 2001; Grant *et al.,* 1993) and in adherence and colonization of *C. jejuni* (Mcsweegan *et al.,* 1986) and *Helicobacter pylori* (Eaton *et al.,* 1996). In *C. difficile,* nonflagellated strains showed a tenfold reduced adherence to tissue in the mouse cecum than flagellated strains belonging to the same serogroup, suggesting that the flagella might have some role in adherence (Tasteyre *et al.,* 2001a; Tasteyre *et al.,* 2001b). Studies on sera from convalescent patients have shown that antibody responses against both FliC and FliD are generated, implicating them as important virulence factors during the course of CDAD (Péchiné *et al.,* 2005a; Péchiné *et al.,* 2005b). Genetic studies demonstrated that inactivation of either the flagellin (*fliC)* or flagellar cap (*fliD)* genes results in complete loss of flagella

as well as motility of *C. difficile* (Dingle *et al.,* 2011). Strikingly, both *C. difficile fliC* and *fliD* mutant strains exhibited higher adherence to enterocyte-like Caco-2 cells and more virulence in hamsters than the wild-type *C. difficile* strain delta-*ermB* (Dingle *et al.,* 2011). Indeed, when *C. difficile* vegetative cells enter in contact with Caco-2 cells, microarray and real-time PCR results suggest that there is a downregulation of the genes involved in flagellar assembly (i.e., *fliH* and *flgG*) (Janvilisri *et al.,* 2010). These results suggest that in *C. difficile* the flagellum is adequately regulated to aid with motility when in the colonic lumen and with adherence when entering in contact with Caco-2 cells (Vedantam *et al.,* 2012).

In accordance with this, Aubry and co-workers (Aubry *et al.,* 2012) have recently demonstrated that the flagellar regulon of *C. difficile* modulates toxin production and that flagellar mutants have a corresponding change in virulence respect to the parental strain in a hamster model of *C. difficile* infection.

Fbp68

Extracellular matrix (ECM) components are required for multiple cellular functions during development, such as proliferation, migration, tissue-specific gene expression and, ultimately, apoptosis (Timpl and Brown, 1996). Surprisingly, several bacterial pathogens are able to use these proteins for adhesion through adhesins localized in the bacterial cell surface. This is also the case for *C. difficile* cells, which interact with several proteins of the ECM, such as fibrinogen, laminin, fibronectin and collagen I, III, IV and V (Cerquetti *et al.,* 2002), suggesting that *C. difficile* vegetative cell adhesins may take advantage of these molecules, using them as molecular bridges to increase adherence to the colonic epithelium.

Fibronectin is a ubiquitous glycoprotein with a molecular size of 450 kDa found in the body fluids and ECM (Henderson *et al.,* 2011). Fibronectin-binding proteins have been well described in Streptococci and Staphylococci, and act as

ligands for fibronectin increasing their ability to infect host cells (Van Der Flier *et al.,* 1995; Navarre and Schneewind, 1999). Interestingly, a gene (*fbp68*) encoding a putative fibronectin binding protein has been found in all draft-assembled and completed genomes of *C. difficile* sequenced to date (Vedantam *et al.,* 2012). Indeed, *C. difficile* Fbp68 is able to bind to the ECM proteins fibronectin, fibrinogen and, to a lesser extent, to vitronectin (Hennequin *et al.,* 2003). A significant reduction in adherence to Vero cells was observed when *C. difficile* cells were pretreated with anti-Fbp68 antibody (Hennequin *et al.,* 2003). In addition, *C. difficile* cells adhered to a lesser extent to Caco-2 cells pretreated with anti-Fbp68 antibody or in siRNA-transfected Caco-2 cells (Lin *et al.,* 2011). *C. difficile fbp68* cells had a slightly slower cecal colonization and intestinal implantation than *C. difficile* wild-type cells, suggesting that fibronectin is exploited to some extent as a molecular bridge to increase colonization of colonic surfaces by *C. difficile* (Barketi-Klai *et al.,* 2011).

4.3. Proteolytic enzymes

Besides Cwp84 and Cwp13, other proteolytic enzymes are likely to play a role in early stages of colonization. An early study showed the production of hyaluronidase, collagenase and chondroitin-4-sulfatase by *C. difficile* (Seddon *et al.,* 1990). A later study examined various *C. difficile* isolates and demonstrated that the supernatant of bacterial cultures contained enzymes that are able to degrade gelatin, collagen and azocoll (Poilane *et al.,* 1998). However, the role of these enzymes in pathogenesis is unclear. They might act as auxiliary virulence factors that facilitate infiltration of *C. difficile* cells and toxins to the colonic mucosa.

Also, proteolytic enzymes may play a role in releasing suitable substrates from available protein sources for metabolism (Seddon and Borriello, 1992; Poilane *et al.,* 1998; Maeda, 1996).

AIM OF THE WORK

To understand the mechanisms of host adaptation, establishment of the disease and transmission from a host to another, it is of outstanding importance to identify and characterize all the factors that are involved in the interaction between the pathogen and its host. So far, most efforts have been focused on characterization of clostridial toxins (toxin A and B, binary toxin), which are recognized as major virulence factors (Kuehne *et al.*, 2010; Lyras *et al.*, 2009; Voth and Ballard, 2005; Schwan *et al.*, 2009). However it is evident that other factors, either exposed on the bacterial surface or released by the bacterium in the extracellular environment, are needed for bacterial fitness and establishment/transmission of disease.

In recent years, the sequencing of several *C. difficile* genomes (Sebaihia *et al.*, 2006; He *et al.*, 2013) and the spread of molecular tools that allow *C. difficile* genetic studies (Cartman *et al.*, 2012; Kuehne *et al.*, 2011) have led to the identification of a certain number of surface/secreted proteins and few of them have been associated to pathogenesis through *in vitro* and *in vivo* studies (Sarker and Paredes-Sabja, 2012). However, a systematic identification of *C. difficile* secreted proteins, which are either exposed on the cell surface or released in the environment, has been never carried out with success. Previous studies have tried to characterize the extracellular *C. difficile* protein profile through proteomic approaches, but in all cases the precise cellular localization of the proteins identified was unclear due to contamination with cytoplasmic proteins (Boetzkes *et al.*, 2012; Wright *et al.*, 2005; Mukherjee *et al.*, 2002).

In order to identify novel factors that could play a role in the host-pathogen interaction, first part of this work was aimed at comprehensively identifying surface-exposed and exported proteins of *C. difficile*. This was assessed taking advantage of a proteomic strategy. Two clinical isolates were analyzed: 630,

isolated from a case of pseudomembranous colitis (Wüst *et al.*, 1982), and R20291, an important hypervirulent strain belonging to the BI/NAP1/027 category (Loo *et al.*, 2005; McDonald *et al.*, 2005).

Second part of this work was aimed at characterizing, among the identified surface or secreted proteins, novel factors likely involved in pathogenesis.

MATERIALS AND METHODS

1. Bacterial strains and culture conditions

Clostridium difficile

C. difficile clinical isolates 630 and R20291 were used in this study. Bacteria were grown in liquid or solid media at 37°C under anaerobic conditions (80% N_2, 10% CO_2, 10% H_2) in a Don Whitley workstation (Yorkshire, United Kingdom).

For proteomic analysis, bacteria were cultured in chemically defined medium (CDMM) prepared as described by Karasawa (Karasawa *et al.*, 1995) or in rich medium (BHIS or TYM). Bacteria from glycerol stocks were grown overnight on BHIS plates. One single colony was inoculated in 5 ml of the desired medium and grown overnight to stationary phase. The resulting culture was diluted in 5 ml of the same medium to a starting OD_{600} of 0.05 and grown to mid-exponential phase (0.4-0.8). The culture was diluted again to a starting OD_{600} of 0.05 in 50 ml of the same medium and grown again up to mid-exponential phase (0.4-0.8).

For the second selection in the genetic manipulation of *C. difficile*, a different minimal medium was used (CDM), prepared according to Cartman and co-workers (Cartman *et al.*, 2012). For immunoblotting analysis, bacteria were grown in TYM broth. For phenotypical characterization of *630/CD630_36690*, bacteria were grown in BHIS broth.

Escherichia coli

E. coli strain DH5alpha (Invitrogen) was used for cloning, T7 Express (New England Biolabs) was used for recombinant proteins expression and purification. *E. coli* was cultured in liquid or solid Luria-Bertani (LB) medium supplemented with suitable antibiotics. To induce expression of recombinant

proteins in T7 Express, liquid medium EnPresso (Biosilta) containing 1 mM IPTG was used.

Compositions of all the culture media used in this study are reported in Appendix 1.

2. Proteomic analysis of *C. difficile* cell surface and culture supernatants

Analysis of *C. difficile* surface

To analyze the surface protein profile, after reaching the exponential phase of growth 50 ml cultures were centrifuged at 3200 *g* for 10 min at 4°C and the medium discarded. A wash step in PBS, pH 7.4 was performed. Then the bacterial pellet was resuspended in digestion buffer and the reaction incubated at 37°C. Several growth media, digestion buffer compositions and times of digestion were tested, trying to find a condition in which no bacterial lysis was detectable. In particular, we changed (i) growth medium (CDMM defined medium, BHIS and TYM rich media); (ii) osmotic conditions (to 750 µl TrisHcl 50 mM, pH 7.5 an osmotic agent was added, either sucrose, in a concentration ranging from 0 to 33%, or NaCl, in a concentration ranging from 0 to 1 M); (iii) trypsin concentration (from 0 to 50 µg); (iv) time of digestion (from 15 min to 16 h).

After digestion, bacteria were pelleted by centrifugation at 16000 *g* for 10 min. The supernatant was filtered through a 0.22 µm filter. To obtain a peptide mixture suitable for HPLC separation, proteins contained in the pellet were digested in small peptides. 0.1% Rapigest (WatersTM, MA, USA) and 5mM DTT were added and the sample was boiled for 10 min. After cooling, pH was checked and eventually ammonium bicarbonate crystals were added to obtain a pH between 7.5 and 8.5. 1 µg of Trypsin was added and the digestion reaction was incubated overnight at 37°C. Digested samples were desalted for

LC/MS/MS analysis by reverse-phased chromatography, using OASIS Cartridges HLB 1cc (WatersTM, MA).

Analysis of *C. difficile* culture supernatants

To analyze the protein content of culture supernatants, after reaching the desired phase of growth in CDMM medium 50 ml cultures were centrifuged at 3200 *g* for 10 min at 4°C and the medium transferred to a clean tube. The supernatant was filtered through a 0.22 µm filter in order to remove any remaining bacteria, and EDTA-free Complete protease inhibitor cocktail (Roche) plus 5 mM EDTA was added. Proteins were precipitated by addition of 10% TCA, 0.04% sodium deoxycholate and incubation for 3 h in ice, followed by centrifugation at 37000 *g* for 20 min at 4°C. The pellet was resuspended in 10% TCA and centrifuged as before. Pellets were washed three times with decreasing amounts of cold absolute ethanol (half volume, 1 ml, 200 µl), then dried in a vacuum concentrator and suspended in 50 mM ammonium bicarbonate; the pH was adjusted to 7.5-8.5 by adding ammonium bicarbonate crystals directly into the tube.

To obtain a peptide mixture suitable for HPLC separation, proteins contained in the pellet were digested in small peptides. 0.1% Rapigest (WatersTM, MA, USA) and 5mM DTT were added and the sample was boiled for 10 min at 100 °C. After cooling, 1 µg of Trypsin was added and the digestion reaction was incubated overnight at 37°C. Digested samples were desalted for LC/MS/MS analysis by reverse-phased chromatography, using OASIS Cartridges HLB 1cc (WatersTM, MA).

Peptide separation and mass spectrometry analysis (LC/MS/MS)

Peptides were separated by nano-LC on a NanoAcquity UPLC system (Waters) connected to a Q-ToF Premier Electro Spray Ionization (ESI) mass spectrometer equipped with a nanospray source (Waters). Samples were loaded

onto a NanoAcquity 1.7 μm BEH130 C_{18} column (75 μm X 25 mm, Waters) through a NanoAcquity 5 μm Symmetry® C_{18} trap column (180 μm X 20 mm, Waters). Peptides were eluted with a 120-min gradient of 2–40% of 98% acetonitrile, 0.1% formic acid solution at a flow rate of 250 nl/min.

The eluted peptides were subjected to an automated data-dependent acquisition using the MassLynx software, version 4.1 (Waters), in which a MS survey scan was used to automatically select multicharged peptides over the *m/z* ratio range of 300–2000 for further MS/MS fragmentation. Up to five different components were subjected to MS/MS fragmentation at the same time. After data acquisition, individual MS/MS spectra were combined, smoothed, and centroided using ProteinLynx, version 3.5 (Waters), to obtain the peak list file. The Mascot Daemon application (MatrixScience Ltd., London, UK) was used for the automatic submission of data files to a version of MASCOT (version 2.2.1) running on a local server.

Protein identification was achieved by searching in a locally curated database combining protein sequence data derived from the *Clostridium difficile* section of the NCBInr database, the total number of sequences and residues being 57275 and 17440799, respectively. The MASCOT search parameters were set to (i) 4 as number of allowed missed cleavages for Trypsin digestion, (ii) methionine oxidation and glutamine and asparagine deamidation as variable modifications, (iii) 0.3 Da as peptide tolerance, and (iv) 0.3 Da as MS/MS tolerance. Only significant hits were considered, as defined by the MASCOT scoring and probability system. The score thresholds for acceptance of peptide identification were ≥48 for trypsin digestion.

3. Bioinformatic prediction of protein subcellular localization

All the nucleotide and aminoacid sequences of proteins were obtained from the NCBI database (http://www.ncbi.nlm.nih.gov/).

The presence of a signal peptide was predicted using the online software SignalP 4.0 (http://www.cbs.dtu.dk/services/SignalP/). Subcellular localization of proteins identified in culture supernatants was predicted using the online software PSORTb 3.0 (http://www.psort.org/psortb/). The presence of a putative prokaryotic lipoprotein motif in CD630_36690 was predicted using the online software PROSITE (http://prosite.expasy.org/).

Aminoacid alignment of sequences of interest on the global protein database was carried out using blastp at http://blast.ncbi.nlm.nih.gov/Blast.cgi.

For alignment of Zmp1 ATLF domain with Anthrax Lethal Factor C-terminal ATLF domain, the ClustalW2 software was used (http://pbil.univ-lyon1.fr/).

4. Preparation of *C. difficile* and culture supernatant fractions

To prepare a total cell extract (TE), strain 630 was grown in TYM to OD_{600} 1.3. Cells were harvested by centrifugation and the pellet was washed and resuspended in PBS. Cell lysis was carried out by freeze-thawing.

In order to isolate the "mutanolysin extract" (ME), containing the S-layer proteins together with other proteins that are present within the cell wall, bacteria were grown in 20 ml of TYM to OD_{600} 1.3. Cells were separated by centrifugation at 3200 *g* and washed once in PBS and once in Tris-sucrose buffer (10 mM Tris-HCl pH 6.9, 10 mM $MgCl_2$, 0.5 M sucrose). The pellet was then incubated in 2 ml of digestion buffer (Tris-sucrose buffer with 250 µg/ml mutanolysin and protease inhibitors) for 2 h at 37°C with gentle agitation. The reaction was centrifuged to separate the supernatant containing the cell wall and S-layer fractions, from the pellet containing the protoplast fraction (P). Protoplasts were resuspended in PBS and lysed by freeze-thawing.

The S-layer fraction (SL), containing only proteins associated to the S-layer, was prepared as described previously (Calabi *et al.*, 2001). Bacteria were grown in 50 ml of TYM broth to OD_{600} 1.3. Cells were separated from the medium by

centrifugation at 3500 *g* for 10 min at RT, washed in PBS and finally incubated in 0.5 ml of 0.2 M HCl, pH 2.2 for 20 min at RT with gentle agitation, in the presence of protease inhibitors. The bacterial suspension was centrifuged at maximum speed at 4°C for 10 min. The supernatant containing the S-layer proteins was removed and the pH neutralized by addition of 2 M Tris base.

To prepare supernatant fractions, bacteria of strain 630 were grown in 100 ml TYM to OD_{600} 1.3 and centrifuged at 3200 *g* to separate the culture supernatant (total supernatant, TS). Half of the total supernatant was ultracentrifuged at 150000 *g* for 16 h at 4°C to obtain two fractions (ultracentrifugation pellet, UP, and ultracentrifugation supernatant, US). To prepare total supernatant and ultracentrifugation supernatant fractions for immunoblotting analysis, proteins were precipitated by addition of TCA, as described for LC/MS/MS analysis, and then resuspended in PBS. The ultracentrifugation pellet was immediately resuspended in PBS.

5. Production of antibodies

Primary antibodies used for immunoblotting analysis of cell and supernatant fractions are polyclonal mouse antisera raised against recombinant his-tagged proteins.

To obtain a serum recognizing all the proteins released by *C. difficile* in culture supernatant, *C. difficile* 630 was grown in CDMM to OD_{600} 0.5. Cultures were centrifuged and culture supernatants filtered through 0.22 µm filters. Proteins present in this fraction were concentrated using Vivaspin centrifugal concentrators MWCO 5000 Da (Sigma). The resulting protein concentration was estimated with the bicinchoninic acid assay (Pierce, Rockford, IL, USA). Eight mice were immunized three times with 20 µg of protein mix per dose. The resulting sera were pooled and used in immunoblotting analysis.

6. Immunoblotting analysis

For immunoblotting analysis of whole-cell lysates, cell fractions, supernatant fractions and purified recombinant proteins, the following amounts of cell and supernatant fractions were loaded:

- TE: 10 µl of total extract 10x concentrated respect to the starting culture
- P: 10 µl of protoplast extract 20x concentrated respect to the starting culture
- ME: 10 µl of mutanolysin extract 10x concentrated respect to the starting culture
- SL: 10 µl of S-layer extract 100x concentrated respect to the starting culture
- TS, US and UP: 10 µl of total supernatant 10x concentrated respect to the starting culture for CD630_18300, CD630_02370 and CD630_23650; 10 µl of total supernatant 250x concentrated respect to the starting culture for CD630_36690
- 10 ng of recombinant protein were loaded as a control

For immunoblotting analysis of recombinant proteins with serum recognizing concentrated culture supernatants, the following amounts of recombinant proteins were loaded:

- 200 ng of CD630_01830, CD630_02370, CD630_14690, CD630_21770, CD630_23650, CD630_27950, CD630_36690
- 2 ng of CD630_05140

All the samples were mixed with NuPage LDS Sample buffer (Invitrogen) and Sample Reducing Agent (Invitrogen) and heated for 10 min at 100°C prior to loading.

Samples were separated by SDS-PAGE on a NuPAGE gel (Invitrogen) and transferred onto nitrocellulose membranes using an iBlot Dry Blotting System (Invitrogen). Membranes were blocked 1 h at RT by agitation in blocking

solution (10% skimmed milk, 0.05% Tween-20, in PBS). All the primary antibodies used were diluted 1:1000 in 1% skimmed milk, 0.05% Tween-20 in PBS and incubated 90 min at 37°C. After washing in 1% milk solution, the membranes were incubated in 1% milk solution in a 1:5000 dilution of peroxidase-conjugated goat anti-mouse serum (Dako) at RT for 45 min. After washing in PBS, 0.05% Tween-20, the resulting signal was detected using the Super Signal West Pico chemiluminescent substrate (Pierce) following the manufacturer's instructions.

7. Site-directed mutagenesis

The pET15-TEV expression vector containing the gene fragment encoding residues 27–220 of Zmp1 was used as a template for site-directed mutagenesis. This construct allows expression of a phusion protein constituted by an N-terminal His6-tag, a cleavage site for TEV protease and a Zmp1 fragment corresponding to the full length protein lacking the predicted signal peptide.

The E143A mutant was generated by site-directed mutagenesis using the polymerase incomplete primer extension (PIPE) system (Klock *et al.*, 2009). Primers E143Af and E143Ar (Appendix 2) were used to substitute the GAA codon (glutamate) with the GCA codon (alanine).

The H146A mutant was generated by site-directed mutagenesis using the GeneArt® Site-Directed Mutagenesis system (Invitrogen). Primers H146Af and H146Ar (Appendix 2) were used to substitute the CAT codon (histidine) with the GCT codon (alanine). Polymerase chain reaction was carried out following manufacturer’s instructions.

The resulting plasmids were used to transform the T7 Express *E. coli* strain (New England Biolabs).

8. Protein expression and purification

Protein expression was performed in T7 Express *E. coli* cells (New England Biolabs). To produce the His-tagged Zmp1 (CD630_28300) wild type or mutants, the cells were cultured overnight at 30°C in EnPresso medium (Biosilta-Oulu, Finland) following the manufacturer's instructions. Protein expression was induced with 1 mM IPTG for 8 h at 30°C. The harvested cells were lysed by sonication in binding buffer (20 mM Tris-HCl, 300 mM NaCl, 10 mM imidazole, pH 8.0) and centrifuged. Tag-less proteins were obtained through a two-step purification protocol (Fig. 1). The supernatant was loaded onto a PD-10 gravity flow empty column (GE Healthcare, NJ, USA) packed with 2 ml of Ni-NTA FF resin (Qiagen GmbH, Hilden, Germany) equilibrated with the binding buffer. The protein was eluted with buffer containing 300 mM imidazole. Buffer was then exchanged in 50 mM TrisHCl, 0.5 M EDTA, 1 mM DTT, pH 8.0 (TEV buffer) by size exclusion chromatography (PD-10 desalting columns, GE Healthcare). Then, cleavage of the His-tag with the tobacco etch virus (TEV) protease (Invitrogen, 5 ng TEV for each µg of recombinant protein) was carried out at RT overnight. After digestion, TEV buffer was substituted with binding buffer and a second "subtractive" purification was performed on the Ni-NTA column. This step allows purification of the protein in the flow-through while cut His-tag, TEV protease, uncut protein and contaminants are retained inside the column. Protein concentrations were estimated with the BCA assay (Pierce, Rockford, IL, USA). The final purity quality of the proteins was checked by SDS-PAGE.

To perform activity assays, the purified proteins were incubated with 20 mM EDTA overnight at 4°C to remove any bound divalent cation and buffer was exchanged with Tris buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.6) by size exclusion chromatography. Different metallated forms of Zmp1 were obtained by addition of $ZnCl_2$, $CuCl_2$ or $NiCl_2$ solutions up to the desired final concentration.

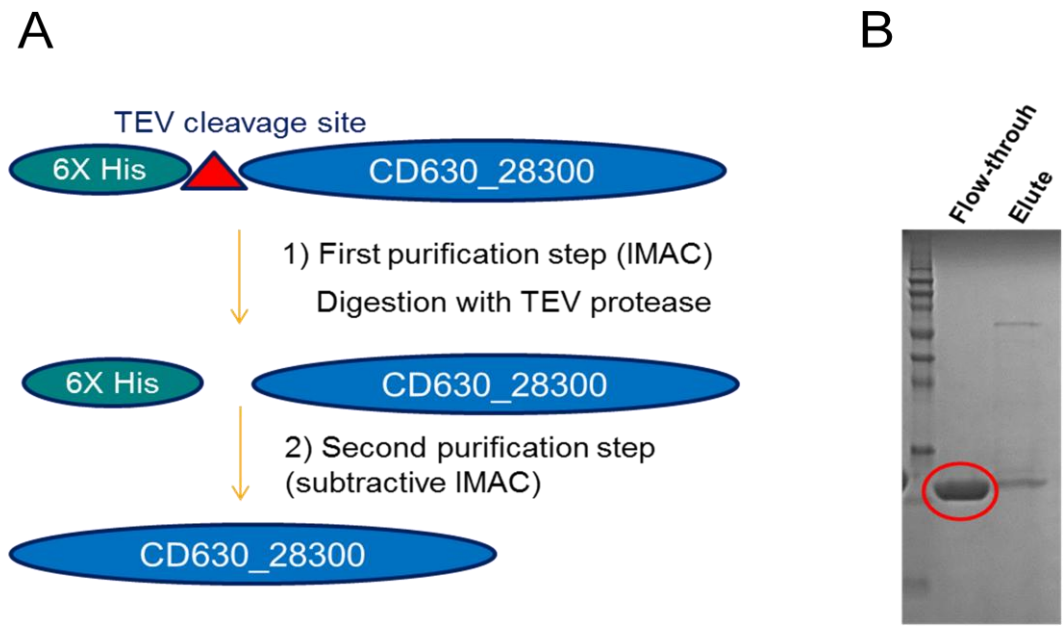


Figure 1. Scheme representing the two-step procedure used for purification of tag-less recombinant proteins. (A) The 6His-tagged protein was purified by IMAC. Then the 6His tag was removed by incubation with TEV protease and a second "subtractive" IMAC was carried out. **(B)** During subtractive IMAC, the tagless protein is obtained in the flow-through, while cut His-tag, TEV protease, uncut protein and contaminants are retained inside the column.

9. Collagenase/gelatinase fluorimetric assay

Proteolytic activity was assayed using the Molecular Probes EnzChek Gelatinase/Collagenase Assay Kit (Invitrogen, USA) with DQ™ Gelatin from pig skin (E-12054) as a fluorescein-conjugate substrate. Analysis was performed using a 96-well plate approach and all assays were done in triplicate. Each well was set up with 100 µl reaction volume containing 100 µg/ml gelatin and 800 µg/ml of recombinant protein.

As a positive control, 0.2 u/ml of control enzyme (purified collagenase type IV from *Clostridium histolyticum*) were used. Reactions were carried out in Tris buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.6) in the absence or presence of 0.5 mM $ZnCl_2$, $CuCl_2$ or $NiCl_2$. Reactions were allowed to proceed for up to 13 h at 37°C. Readings were taken every 10 min using a 96-well fluorescent plate reader (Infinite M200 spectrophotometer microplate reader, Tecan) with

excitation/emission at 480/525 nm, respectively. All values were corrected against the fluorescent curve of a reaction containing everything except the protease.

10. *In vitro* fibronectin cleavage assay

To test for cleavage of fibronectin, 170 µg/ml of fibronectin from human plasma (SIGMA) were incubated at 37°C for 24 h in Tris buffer (50mM TrisHCl, 150mM NaCl, pH 7.6) in absence or presence of an equal amount of Zmp1 and in absence or presence of 0.5 mM $ZnCl_2$, $NiCl_2$ or $CuCl_2$. Reactions were analyzed on a 3-8% NuPAGE gel (Invitrogen) by silver staining with SilverQuest (Invitrogen).

11. Proteolytic activity of Zmp1 protein on human fibroblasts

IMR-90 cells (Human fibroblasts, ATCC) were maintained in Eagle's Minimum Essential Medium (EMEM, ATCC) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Gibco) and antibiotics at 37°C in 5% CO_2. To study Zmp1 activity, cells were seeded on 8-well chamber slides coated with collagen I (2.5×10^4 per well) (BD BioCoat) and grown for 2 days. Cells were then incubated with 250 µg/ml of recombinant Zmp1 protein or mutant Zmp1 proteins and 0.1 mM $ZnCl_2$ for 16 h at 37°C, in EMEM without FBS to avoid the interference of plasma fibronectin. After washings, cells were fixed with 3.7% paraformaldehyde. Samples were washed extensively and incubated with rabbit anti-fibronectin antibodies (Sigma-Aldrich) for 1 h at RT. After multiple washings, samples were incubated with Alexa Fluor 568 goat anti-rabbit IgG. Glass coverslips were mounted with ProLong® Gold antifade reagent with DAPI and analyzed with a Zeiss LSM710 confocal microscope.

12. Differential scanning fluorimetry (DSF)

The shifts in the melting temperature of Zmp1 wild type, E143A and H146A as a function of Zn^{2+} concentration were determined with DSF analysis, following the previously reported protocol (Niesen *et al.*, 2007). In each well, apo protein was incubated at a concentration of 20 μM in 20 μl Tris buffer, in presence of $ZnCl_2$ in a final concentration ranging from 0 to 0.5 mM. The plate was subjected to a temperature gradient scan (25–95 °C) in a Real Time PCR machine (Agilent Technologies, Santa Clara, CA, USA). The fluorescence intensity was measured by use of fluorescence Sypro Orange dye (×5 final concentration) at different temperatures with excitation/emission wavelengths of 490 and 575 nm, respectively.

13. Generation of *C. difficile 630/*Δ*CD630_36690* deletion mutant

Generation of the knock-out plasmid

The 630/Δ*CD630_36690* strain was generated in the 630 background employing the allele exchange strategy described by Cartman and co-workers (Cartman *et al.*, 2012) (Fig. 3). Two allele exchange cassettes were cloned into the pMTL-SC7315 vector (Fig. 2 and Fig. 3). Approximately 500 bp fragments of the flanking regions of the target gene were amplified by PCR from 630 genomic DNA (isolated using a DNeasy Blood and Tissue kit, Qiagen). Amplification reaction was carried out with Phusion High-Fidelity DNA polymerase (New England BioLabs) using the primers couples U3669f/U3669r for amplification of the upstream cassette and D3669f/D3669r for amplification of the downstream cassette (Appendix 2). Primers have ends containing the appropriate restriction sites for cloning (U3669f and D3669r carry a PmeI site, U3669r and D3669f an EcorI site). Thermocycling conditions were 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, an appropriate annealing

temperature for 30 s, 72°C for 1 min/kb, and a final extension of 72°C for 10 min. The upstream and downstream fragments were cloned into the PmeI site of pMTL-SC7315 by one-step ligation with T4 DNA ligase (Rapid DNA Ligation kit, Roche) to generate an in-frame deletion in the CD630_36690 open reading frame (deletion of nucleotides 4 to 558). The reaction was transformed in the *E. coli* cloning strain DH5alpha (Invitrogen) by heat shock. Selection of transformants was carried out on LB solid medium supplemented with 25 µg/ml chloramphenicol. Several colonies were picked and inoculated in liquid LB medium supplemented with 12.5 µg/ml chloramphenicol. Using standard plasmid extraction methods, plasmid was purified from each of the overnight cultures and screened by restriction analysis with PmeI and by PCR using primers couple SC7-F1/SC7-R1 (Appendix 2). Then the DNA sequence was checked by Sanger sequencing using primers SC7-F and SC7-R (Appendix 2). Restriction endonucleases were obtained from New England BioLabs and used according to the manufacturer's instructions.

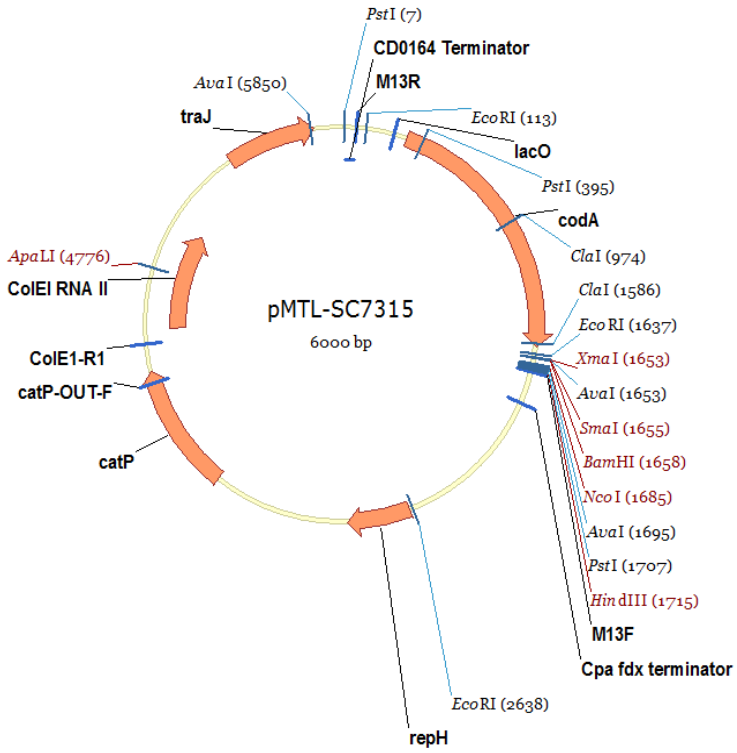


Figure 2. Map of pMTL-SC7315 vector for allele exchange.

Transfer of knock-out plasmid into the *C.difficile* recipient by conjugation

A knockout cassette plasmid verified by sequencing was re-transformed into electrocompetent *E.coli* conjugation donor strain CA434, carrying out selection on LB plates supplemented with 20 µg/ml chloramphenicol. Recipient strain *C. difficile* 630 and donor strain *E. coli* CA434/pMTL-SC7315 were grown overnight in BHIS and in LB/12.5 µg/ml chloramphenicol, respectively. 1 ml of the stationary overnight culture of *E. coli* donor strain was pelleted, washed and the cells were re-suspended in 0.5 ml PBS. The conjugal donor pellet was then resuspended in 200 µl of stationary overnight culture of conjugal recipient cells and the entire conjugation mixture was pipetted onto a single BHIS non-selective plate in discrete drops. After 24 h of incubation under anaerobic conditions (to allow conjugal transfer of the knockout plasmid from the *E. coli* donor to the *C. difficile* recipient) all growth was harvested into 500 µl PBS and plated onto BHIS plates containing D-cycloserine (250 µg/ml), cefoxitin (8 µg/ml) and thiamphenicol (115 µg/ml) to select for transconjugants (i.e. *C. difficile* cells harbouring the plasmid). Plates were incubated at 37°C under anaerobic conditions for 24-72 h.

Isolation of single cross-over clones

Transconjugant colonies were restreaked and incubated 24-72 h at 37°C. Single cross-over clones were identified as faster growing colonies in amongst slower growing transconjugant colonies. Single cross-over clones were purified by picking and restreaking and then confirmed by PCR analysis carried out with Platinum Taq (Invitrogen) in Failsafe PCR buffer E (Epicentre) using primers pairs ext3669f/SC7-F1, int3669f/SC7-F1, SC7-R1/int3669r, SC7-R1/ext3669r (Appendix 2) (Fig. 3).

Isolation of double cross-over clones

To allow rare second recombination events to occur, single-crossover clones were restreaked onto nonselective BHIS medium and incubated for 96 h. All growth was harvested in 500 µl PBS, serial dilutions were made (from 10^{-1} to 10^{-6}), and 100 µl of each dilution were plated onto CDM supplemented with 50 µg/ml 5-fluorocytosine (FC). After 48 h of incubation, FC-resistant clones were patch plated onto BHIS supplemented with cycloserine, cefoxitin, and thiamphenicol to screen for plasmid loss. Fluorocytosine-resistant thiamphenicol-sensitive clones were analyzed by PCR to distinguish double-crossover recombinant clones from wild-type revertant clones, using primers ext3669f/ext3669r (Appendix 2). Sanger sequencing was used to confirm the expected genotype (Fig. 3).

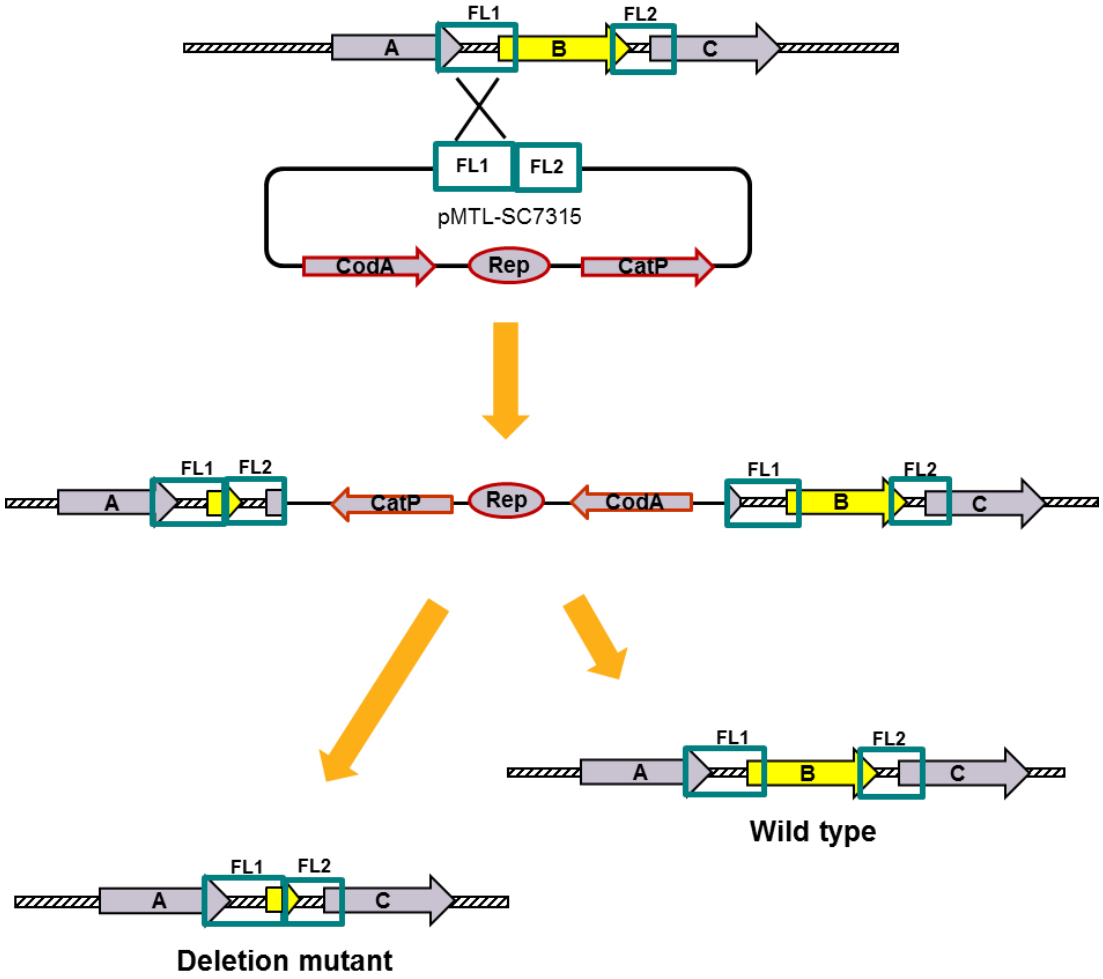


Figure 3. Scheme representing the allele exchange strategy used to generate deletion mutants in *C. difficile*.

14. Vital counts of *C. difficile* spores

Sporulating cultures of *C. difficile* were prepared in BHIS broth. At different times, samples (500 µl) were treated with 50% ethanol for 1 h inside the anaerobic chamber to selectively kill the vegetative cells (Borriello and Honour, 1981). Samples were then washed once with 1 ml PBS, serially diluted in PBS, and plated onto BHIS agar supplemented with 0.1% taurocholate (Sigma, United Kingdom). CFU were enumerated after incubation for 24 h inside the

anaerobic chamber. As a control, total counts of vegetative cells and spores were determined by plating dilutions of untreated culture on the same medium.

15. Confocal immunofluorescence microscopy

C. difficile 630 was grown in BHIS broth to OD_{600} 1.3. Cells were pelleted by centrifugation at 3200 *g* for 10 min at 4°C, washed in PBS and then fixed in double volume of PBS/4% formaldehyde for 20 min at RT. After washing in PBS, cells were resuspended in the equal volume of PBS and 20 µl of bacterial solution were spotted on a glass slide pretreated with a 10% V/V solution of 0.1% poly-lysine (Sigma). Blocking was carried out with 3% BSA in PBS for 20 min at RT. Primary antibody was diluted 1:500 in 1% BSA/PBS for 1 h at RT. After 3 washes in PBS, slides were incubated with Alexa Fluor® 568 Goat Anti-Mouse IgG (H+L) (Invitrogen) diluted 1:400 in 1% BSA/PBS for 30 min at RT. Slides were washed 3 times in PBS and mounted with ProLong Gold Antifade Reagent with DAPI (Invitrogen). Slides were analyzed with Zeiss LSM710 confocal microscope.

16. Electron microscopy

Droplets of OMV suspension (10 µl) were placed on formvar-carbon coated grids and allowed to adsorb for 60 s. Excess liquid was removed gently touching the filter paper. The adsorbed specimen was then fixed for 2 min at room temperature floating on a drop of 4 % paraformaldehyde plus 0.05% glutaraldehyde in PBS, pH 7.2. Negative-staining was performed by first washing the specimen grid on a drop of negative stain (2% uranyl acetate in distilled water), blotting and repeating this step once more, this time leaving the specimen grid for 60 s on a new drop of negative stain solution.

Samples were observed at a JEOL 1200 EX II electron microscope. Micrographs were acquired by the Olympus SIS VELETA CCD camera equipped with the iTEM software.

RESULTS

1. Identification of surface and secreted proteins of *C. difficile* by mass spectrometry

1.1. Experimental approach

In order to comprehensively identify proteins of *C. difficile* that are associated to the surface or released in the extracellular environment, two different proteomic approaches were tried (Fig. 1).

The first approach, known as "surface shaving", has successfully been used to characterize the surface protein profile of several bacterial species (Bensi *et al.*, 2012; Doro *et al.*, 2009; Rodriguez-Ortega *et al.*, 2006; Berlanda Scorza *et al.*, 2008). After growth in liquid medium, bacteria are treated with trypsin, a protease that should act on the bacterial surface releasing into the medium a mixture of peptides deriving from surface proteins. These peptides are then recovered, separated by liquid chromatography and analyzed by mass spectrometry (LC/MS/MS).

The second approach, based on a similar analysis of culture supernatants (Desvaux *et al.*, 2010; Bensi *et al.*, 2012), is aimed at identifying proteins that are released by bacteria in the medium during growth. Proteins contained in culture medium are precipitated and then digested with trypsin in order to produce a mixture of peptides that can be analyzed by LC/MS/MS.

For this analysis, two clinical isolates were selected, 630 and R20291. 630 is a reference strain for *C. difficile*, as it is the first for which genome sequencing was completed (Sebaihia *et al.*, 2006). It was isolated from a case of pseudomembranous colitis (Wüst *et al.*, 1982). R20291 is an important epidemic strain belonging to the category BI/NAP1/027, which caused an

increase in rates and severity of CDAD in North America and Europe (Loo *et al.*, 2005; McDonald *et al.*, 2005).

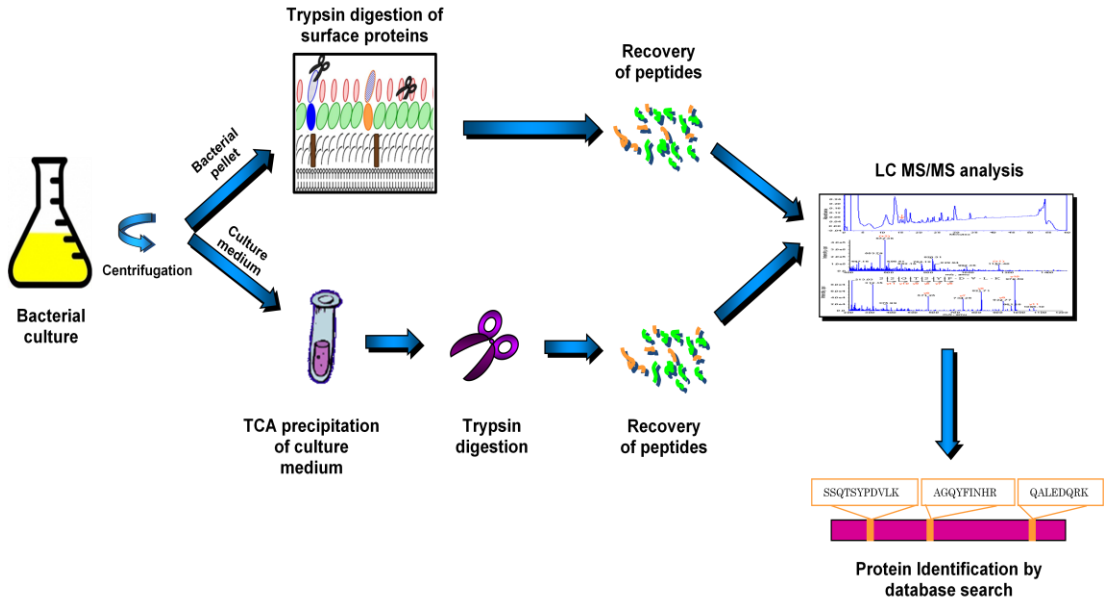


Figure 1. Scheme representing the two proteomic approaches used to identify surface and released proteins of *C. difficile*.

1.2. Set up of growth conditions for proteomic analysis

For proteomic analyses, bacteria were grown in the desired medium to the exponential phase of growth (OD_{600} 0.4-0.8). Growth curves were performed for both strains 630 and R20291 in all the media in order to identify the range of optical density in which the cultures are in exponential phase (Fig. 2A). Both strains showed similar growth profiles, reaching maximum OD_{600} of 1.2-1.3 in chemically defined medium and of 1.8-2.0 in rich media (Fig. 2A). To analyze bacteria and culture media in exponential phase, bacteria were collected in a range of OD_{600} between 0.4 and 0.8. In order to minimize the amount of bacterial cell lysis, and subsequent contamination with cytoplasmic proteins,

bacteria were grown to the exponential phase multiple times, as described in Methods, before collecting supernatant samples for LC/MS/MS analysis.

Quantitation of spores from cultures used for sample preparation revealed a very low percentage of spores (0.01%), suggesting an irrelevant contribution of proteins originating from them (Fig. 2B).

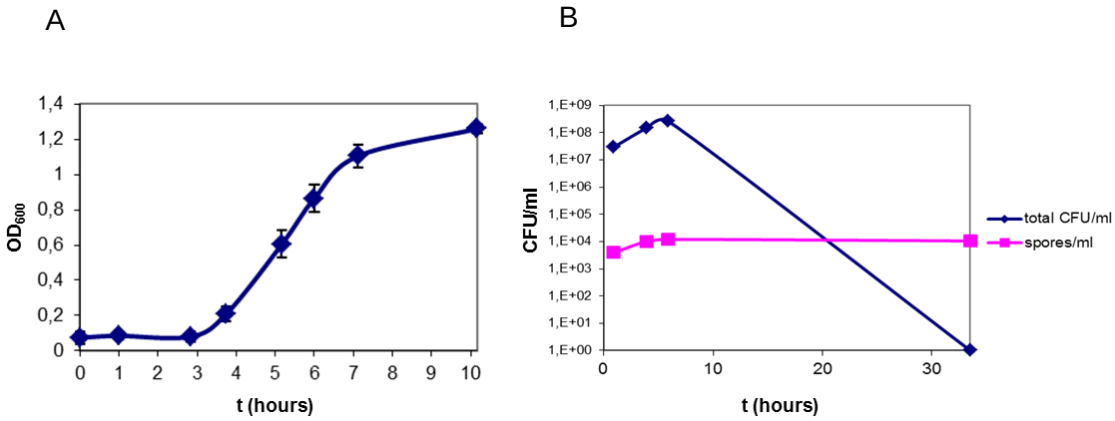


Figure 2. Growth profile and vital counts of vegetative cells and spores for strain 630 in CDMM. (**A**) Growth curve of strain 630 in CDMM (growth curves were performed in triplicate). R20291 strain showed a similar growth profile. (**B**) Vital counts of vegetative cells and spores of strain 630 grown in CDMM.

1.3. Analysis of *C. difficile* cell surface

For surface shaving of *C. difficile*, several experimental conditions were tested (Table 1). However, in all cases a high amount of cytoplasmic proteins was detected in the reaction mixture (data not shown), indicating bacterial lysis. The analysis was carried out on both 630 and R20291 strains, as different bacterial isolates could respond differently to enzymatic treatments due to differences in their surface organization; however with both isolates a similar amount of bacterial lysis was observed. Several growth media (BHIS, TYM, CDMM and BHIS agar) were tested, in an attempt to minimize the physiologic lysis that takes place during bacterial growth. Moreover, many different

conditions were changed for the shaving reaction. On one side, given the importance of an osmotic agent in preventing cell lysis, two osmotic agents (sucrose and NaCl) were tested at different concentrations. On the other side, several concentrations of protease were tried, including low concentrations (0.25 and 2.5 µg/ml), in an attempt to carry out a weaker surface digestion, and high concentrations (65 µg/ml), in an attempt to increase the ratio between the number of surface peptides and the number of cytoplasmic proteins released into the reaction mixture. To understand if trypsin was involved in bacterial lysis (i.e., causing a very rapid digestion, with subsequent disruption of cells), we performed control reactions in digestion buffer containing everything but trypsin; however also in this case we detected a large percentage of cytoplasmic proteins, indicating that other factors than trypsin must be involved in bacterial lysis.

Presence of cytoplasmic proteins in the samples does not allow to assign the identified proteins to the correct subcellular compartment. Therefore, this approach did not allow the identification of *C. difficile* surface proteins.

Strain	Growth medium	Osmotic agent	Trypsin concentration	Time of trypsin incubation
630	TYM	Sucrose (0-33%)	0 µg/ml	15'
R20291	CDMM	NaCl (0-1 M)	0,25 µg/ml	20'
	BHI		2,5 µg/ml	30'
	BHI agar		13 µg/ml	2 h
			65 µg/ml	16 h

Table 1. Experimental conditions tested for surface shaving of *C. difficile*.

1.4. Analysis of *C. difficile* culture supernatants

Analysis of 630 and R20291 culture supernatants led to the identification of a total of 51 proteins for strain 630 (from 6 experiments) and 30 proteins for R20291 (from 3 experiments) (Tables 2 and 3, Fig. 3). There was a considerable overlap (25 proteins) between the proteins identified in the two strains; however 26 proteins were detected only in 630 and 5 only in R20291, although all the coding genes, except CD630_23880 and CDR20291_2278, are present in both genomes with at least 80% identity. We observed extremely low numbers of predicted cytoplasmic proteins, indicating insignificant levels of bacterial cell lysis during growth (Tables 2 and 3).

Interestingly, majority of the proteins identified were putative surface-associated proteins (Fig. 3). We grouped most of these proteins into five major classes, based primarily on the presence of domains responsible for surface association and secondly on the presence of conserved functional domains: CWPs (cell wall proteins), LPXTG-anchored proteins, extracellular hydrolases, transporters and flagellar proteins.

Based on the presence of domains potentially responsible for surface association, we found 14 proteins containing the cell wall binding repeat 2 domain (pfam04122), which is characteristic of the clostridial cell wall proteins (CWPs) (Sebaihia *et al.,* 2006). These include previously studied proteins such as CwpV (Emerson *et al.,* 2009; Reynolds *et al.,* 2011), Cwp84 (Janoir *et al.,* 2007; Kirby *et al.,* 2009), Cwp66 (Waligora *et al.,* 2001), Cwp2 (Calabi *et al.,* 2002b; de la Riva *et al.,* 2011), SlpA (Calabi *et al.,* 2001; Fagan *et al.,* 2009). Interestingly, Cwp24, recently annotated as a pseudogene or gene remnant in strain 630 (Monot *et al.,* 2011), was identified in both 630 and R20291.

We also identified two proteins (CD630_28310, CD630_01830) containing a LPXTG-like motif at the C-terminus, a motif associated with sortase-mediated anchoring to the peptidoglycan (Sebaihia *et al.,* 2006).

A number of cell wall hydrolases with extracellular or membrane localization were detected. Among them, the previously characterized autolysin Acd (Dhalluin *et al.,* 2005). Also, numerous ABC transporters, predicted to be involved in the uptake of sugars, amino acids, peptides, or inorganic ions were identified. Flagellar components were also a major protein family detected in the culture supernatants; notably, whereas the major flagellar subunit (FliC) was present in both strains, other proteins predicted to be part of the basal body, hook and cap structures were identified only in 630. It is possible that R20291 flagella are less easily detached from the cell, as compared with 630 or their components are less expressed in the growth conditions tested in this study.

Six proteins annotated as hypothetical, for which a function has not been predicted yet, were also detected (Table 2, Fig. 3). For most of them, a cellular localization could not be predicted by PSORTb. However, a cleavage site for signal peptidase could be predicted using SignalP for all of them.

We did not detect *C. difficile* toxins A and B in the supernatants. As it is known from the literature, these toxins are produced after several hours of stationary phase and expression is downregulated at the transcriptional level in the presence of glucose (Mukherjee *et al.,* 2002). Therefore it is not surprising that the two *C. difficile* major toxins were not detected in the experimental conditions used .

Thus, using a proteomic approach, we have identified a subset of putative surface proteins that are extracellular to *C. difficile.*

Family	Gene annotation (630)	Locus tag 630	Signal peptide	Manually cured PSORTb output	630 SN	R20291 SN (locus tag)
CWPs	precursor of the S-layer proteins (SlpA)	CD630_27930	yes	cell wall	yes	CDR20291_2682
	cell wall binding protein (Cwp2)	CD630_27910	yes	cell wall	yes	CDR20291_2680
	cell wall binding protein (Cwp11)	CD630_27950	yes	cell wall	yes	CDR20291_2684
	hemagglutinin/adhesin (CwpV)	CD630_05140	yes	cell wall	yes	nd
	cell surface protein penicillin-binding protein (Cwp20)	CD630_14690	yes	cell wall	yes	nd
	pseudogene (Cwp24)	CD630_21930	yes	cell wall	yes	CDR20291_2099
	cell wall-binding protein (Cwp18)	CD630_10470	yes	cell wall	yes	nd
	cell surface protein (Cwp66)	CD630_27890	yes	cell wall	yes	nd
	cell wall binding protein (Cwp12)	CD630_27940	yes	cell wall	yes	CDR20291_2683
	cell wall binding protein (Cwp9)	CD630_27980	yes	cell wall	yes	nd
	cell surface protein (Cwp84)	CD630_27870	yes	cell wall	yes	nd
	cell wall-binding protein (Cwp25)	CD630_08440	yes	cell wall	no	CDR20291_0774
	N-acetylmuramoyl-L-alanine amidase, autolysin (Cwp6)	CD630_27840	yes	cell wall	yes	CDR20291_2672
	N-acetylmuramoyl-L-alanine amidase (Cwp16)	CD630_10350	yes	membrane	yes	CDR20291_0891
LPXTG-like	adhesin	CD630_28310	yes	cell wall	no	CDR20291_2722
	cell wall hydrolase	CD630_01830	yes	extracellular	yes	CDR20291_0184
Extracellular hydrolases	cell-wall hydrolase	CD630_27680	yes	extracellular	yes	CDR20291_2656
	peptidoglycan-binding/hydrolysing protein	CD630_23880	yes	unknown	yes	-
	peptidoglycan-binding/hydrolysing protein	-	no	unknown	-	CDR20291_2278
	hydrolase	CD630_24800	yes	unknown	yes	nd
	cell wall hydrolase; phosphatase-associated protein	CD630_24020	yes	extracellular	yes	nd
	mannosyl-glycoprotein endo-beta-N-acetylglucosamidase (Acd)	CD630_13040	yes	extracellular	yes	CDR20291_1145
	polysaccharide deacetylase	CD630_15220	yes	extracellular	yes	CDR20291_1371
	lytic transglycosylase	CD630_02260	no	membrane*	yes	nd
	LmbE-like deacetylase	CD630_27900	no	membrane**	yes	nd
Transporters	ABC transporter sugar-family extracellular solute-binding protein	CD630_08730	yes	extracellular*	yes	CDR20291_0802
	ABC transporter sugar-family extracellular solute-binding protein	CD630_08760	yes	extracellular*	yes	CDR20291_0805
	ABC transporter iron-family extracellular substrate-binding protein	CD630_16500	yes	extracellular*	yes	CDR20291_1548
	ABC transporter cystine/aminoacid-family extracellular solute-binding protein	CD630_21770	yes	extracellular*	yes	CDR20291_2083
	ABC transporter oligopeptide-family solute-binding protein	CD630_26720	yes	extracellular*	yes	CDR20291_2560
	ABC transporter cystine/aminoacid-family extracellular solute-binding protein	CD630_21740	yes	extracellular*	yes	nd
	ABC transporter nitrate/sulfonate/taurine extracellular solute-binding protein	CD630_23650	yes	extracellular*	yes	CDR20291_2252
	multidrug efflux pump	CD630_24080	yes	membrane*	yes	CDR20291_2298
Flagellar proteins	flagellar hook-associated protein FlgK (or HAP1)	CD630_02310	no	extracellular*	yes	nd
	flagellar hook-associated protein 2 FliD (or HAP2)	CD630_02370	no	flagellar	yes	nd
	flagellin C	CD630_02390	no	extracellular*	yes	CDR20291_0240
	flagellar hook protein FlgE (Distal rod protein)	CD630_02550	no	flagellar	yes	nd
	flagellar basal body rod protein FlgG	CD630_02690	no	flagellar	yes	nd
	flagellar basal-body rod protein FlgB	CD630_02450	no	extracellular*	yes	nd
	flagellar hook-associated protein FlgL (or HAP3)	CD630_02320	no	extracellular	yes	nd
	flagellar basal-body rod protein FlgC	CD630_02460	no	extracellular*	yes	nd
Hypothetical proteins	hypothetical protein	CD630_36690	yes	unknown	yes	CDR20291_3529
	hypothetical protein	CD630_28300	yes	unknown	yes	CDR20291_2721
	hypothetical protein	CD630_21270	yes	membrane	yes	nd
	hypothetical protein	CD630_05490	yes	unknown	yes	CDR20291_0474
	hypothetical protein	CD630_11560	yes	unknown	yes	nd
	hypothetical protein	CD630_22510	yes	unknown	yes	CDR20291_2151
Other predicted surface proteins	pilin protein	CD630_35130	yes	membrane	yes	nd
	Signal peptidase I, S26A family	CD630_13310	no	cell wall	yes	nd
	SH3-domain-containing protein	CD630_11350	yes	extracellular	yes	CDR20291_0971
	calcium-binding adhesion protein	CD630_27970	yes	unknown	yes	nd
Predicted cytoplasmic proteins	translation inhibitor endoribonuclease	CD630_25130	no	cytoplasmic	yes	nd
	enolase (2-phosphoglycerate dehydratase)	CD630_31700	no	cytoplasmic	yes	nd
	glyceraldehyde-3-phosphate dehydrogenase	CD630_31740	no	cytoplasmic	no	CDR20291_3030
	elongation factor Tu	CD630_00580	no	cytoplasmic	yes	CDR20291_0051
	XkdK-like protein	CD630_13630	no	cytoplasmic	no	CDR20291_1206

Table 2. List of proteins detected by LC/MS/MS in *C. difficile* 630 and R20291 culture supernatants

1. Proteins identified have been divided in families based primarily on the presence of motifs for association to the peptidoglycan and secondly on the presence of conserved functional domains.
2. Presence of a signal peptide was predicted with SignalP 4.0 using the 630 aminoacid sequence (except for CDR20291_2278, for which a sequence with aminoacid identity >40% is not present in 630).
3. Cellular localization was predicted with PSORTb 3.0 using the 630 aminoacid sequence (except for CDR20291_2278); the software output is reported, if not indicated with * (manually cured, deduced by similarity) or ** (manually cured, associated to the major S-layer locus).
4. Detection in 630 culture supernatants (6 replicates) is indicated with 'yes'; 'no' indicates that in 630 an a.a. sequence homologous to the R20291 sequence is present, but the protein was not detected; '-' indicates that in 630 an a.a. sequence homologous to the R20291 sequence is not present .
5. Detection in R20291 culture supernatants (3 replicates) is indicated with the locus tag; 'nd' indicates that in R20291 an a.a. sequence homologous to the 630 sequence is present, but the protein was not detected; '-' indicates that in R20291 an a.a. sequence homologous to the 630 sequence is not present.

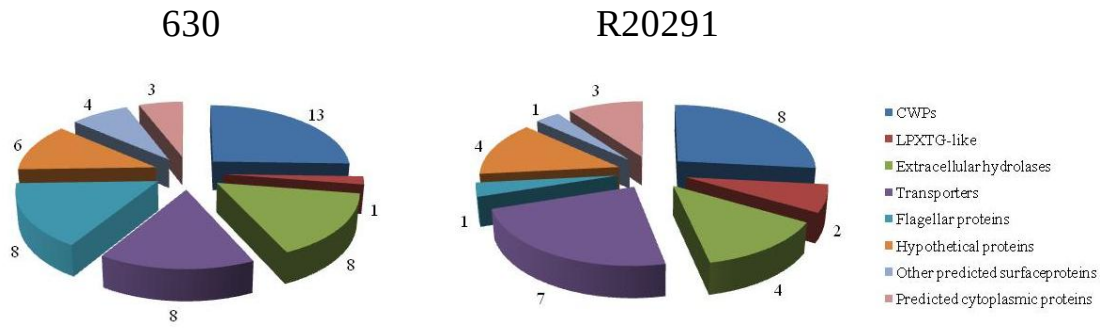


Figure 3. Family distribution of putative surface/released proteins detected in 630 (51 proteins) and R20291 (30 proteins) culture supernatants. Proteins identified have been divided in families based on the presence of motifs for association to the peptidoglycan and on the presence of conserved functional domains.

Family	Gene annotation (630)	Locus tag 630	Presence in CD630 supernatant						Presence in CDR20291 supernatant (locus tag)		
			Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 1	Exp 2	Exp 3
CWPs	precursor of the S-layer proteins (SlpA)	CD630_27930	X	X	X	X	X	X	CDR20291_2682	CDR20291_2682	CDR20291_2682
	cell wall binding protein (Cwp2)	CD630_27910			X			X		CDR20291_2680	CDR20291_2680
	cell wall binding protein (Cwp11)	CD630_27950			X			X			CDR20291_2684
	hemagglutinin/adhesin (CwpV)	CD630_05140		X		X	X	X			
	cell surface protein penicillin-binding protein (Cwp20)	CD630_14690	X	X	X	X	X	X			
	pseudogene (Cwp24)	CD630_21930	X		X	X	X	X	CDR20291_2684	CDR20291_2684	CDR20291_2684
	cell wall-binding protein (Cwp18)	CD630_10470			X						
	cell surface protein (Cwp66)	CD630_27890			X						
	cell wall binding protein (Cwp12)	CD630_27940			X						CDR20291_2683
	cell wall binding protein (Cwp9)	CD630_27980			X						
	cell surface protein (Cwp84)	CD630_27870					X				
	cell wall-binding protein (Cwp25)	CD630_08440							CDR20291_0774		
	N-acetylmuramoyl-L-alanine amidase, autolysin (Cwp6)	CD630_27840	X	X	X	X	X	X	CDR20291_2672	CDR20291_2672	CDR20291_2672
	N-acetylmuramoyl-L-alanine amidase (Cwp16)	CD630_10350	X		X			X			CDR20291_0891
LPXTG-like	adhesin	CD630_28310							CDR20291_2722		
	cell wall hydrolase	CD630_01830	X	X	X	X	X	X	CDR20291_0184	CDR20291_0184	CDR20291_0184
Extracellular hydrolases	cell-wall hydrolase	CD630_27680	X	X	X	X	X	X	CDR20291_2656	CDR20291_2656	CDR20291_2656
	peptidoglycan-binding/hydrolysing protein	CD630_23880	X		X			X			
	peptidoglycan-binding/hydrolysing protein	-							CDR20291_2278	CDR20291_2278	CDR20291_2278
	hydrolase	CD630_24800	X		X	X	X	X			
	cell wall hydrolase; phosphatase-associated protein	CD630_24020	X	X	X						
	mannosyl-glycoprotein endo-beta-N-acetylglucosamidase (Acd)	CD630_13040			X						CDR20291_1145
	polysaccharide deacetylase	CD630_15220			X						CDR20291_1371
	lytic transglycosylase	CD630_02260	X		X						
	LmbE-like deacetylase	CD630_27900	X		X		X	X			
Transporters	ABC transporter sugar-family extracellular solute-binding protein	CD630_08730	X	X	X	X	X	X	CDR20291_0802	CDR20291_0802	CDR20291_0802
	ABC transporter sugar-family extracellular solute-binding protein	CD630_08760			X	X			CDR20291_0805	CDR20291_0805	CDR20291_0805
	ABC transporter iron-family extracellular substrate-binding protein	CD630_16500		X	X	X	X	X	CDR20291_1548	CDR20291_1548	CDR20291_1548
	ABC transporter cystine/aminoacid-family extracellular solute-binding protein	CD630_21770	X	X	X	X	X	X		CDR20291_2083	CDR20291_2083
	ABC transporter oligopeptide-family solute-binding protein	CD630_26720	X	X	X	X	X	X	CDR20291_2560	CDR20291_2560	CDR20291_2560
	ABC transporter cystine/aminoacid-family extracellular solute-binding protein	CD630_21740			X						
	ABC transporter nitrate/sulfonate/taurine extracellular solute-binding protein	CD630_23650	X		X	X	X	X	CDR20291_2252		CDR20291_2252
	multidrug efflux pump	CD630_24080			X				CDR20291_2298		
Flagellar proteins	flagellar hook-associated protein FlgK (or HAP1)	CD630_02310	X		X	X	X	X			
	flagellar hook-associated protein 2 FliD (or HAP2)	CD630_02370	X	X	X	X	X	X			
	flagellin C	CD630_02390	X	X	X	X	X	X	CDR20291_0240	CDR20291_0240	CDR20291_0240
	flagellar hook protein FlgE (Distal rod protein)	CD630_02550	X		X	X	X	X			
	flagellar basal body rod protein FlgG	CD630_02690	X					X			
	flagellar basal-body rod protein FlgB	CD630_02450	X								
	flagellar hook-associated protein FlgL (or HAP3)	CD630_02320	X	X	X		X	X			
	flagellar basal-body rod protein FlgC	CD630_02460	X		X		X	X			
Hypothetical proteins	hypothetical protein	CD630_36690			X			X			CDR20291_3529
	hypothetical protein	CD630_28300	X		X	X	X	X			CDR20291_2721
	hypothetical protein	CD630_21270	X		X		X	X			
	hypothetical protein	CD630_05490	X	X	X		X	X		CDR20291_0474	CDR20291_0474
	hypothetical protein	CD630_11560			X			X			
	hypothetical protein	CD630_22510			X				CDR20291_2151		
Other predicted surface proteins	pilin protein	CD630_35130			X						
	Signal peptidase I, S26A family	CD630_13310			X		X	X			
	SH3-domain-containing protein	CD630_11350			X		X			CDR20291_0971	CDR20291_0971
	calcium-binding adhesion protein	CD630_27970			X			X			
Predicted cytoplasmic proteins	translation inhibitor endoribonuclease	CD630_25130						X			
	enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase)	CD630_31700						X			
	glyceraldehyde-3-phosphate dehydrogenase	CD630_31740									CDR20291_3030
	elongation factor Tu	CD630_00580					X		CDR20291_0051		CDR20291_0051
	XkdK-like protein	CD630_13630									CDR20291_1206

Table 3. List of proteins detected in each experiment in *C. difficile* 630 and R20291 culture supernatants.

2. *C. difficile* 630 produces membrane vesicles during growth *in vitro*, but these are not responsible for the massive release of proteins observed during growth

2.1. Working hypothesis

We hypothesized several mechanisms by which *C. difficile* could release a high number of surface-associated proteins in culture supernatants (Fig. 4). Proteins could be released as single units, not in association to larger complexes, through a passive release (shedding) or through an active cleavage on the bacterial surface. It has already been shown that surface proteases, such as Cwp84 and Cwp13, are involved in *C. difficile* surface remodeling (de la Riva *et al.*, 2011; Kirby *et al.*, 2009). On the other side, proteins could be released in the extracellular environment in association to macromolecular structures, such as cell wall fragments or membrane vesicles (Fig. 4).

Membrane vesicles have been extensively involved in the release of proteins in Gram-negative species (Wai *et al.*, 2003; Kesty & Kuehn, 2004; Bomberger *et al.*, 2009). More recently, similar structures have been observed in supernatants of Gram-positive bacteria (Gurung *et al.*, 2011; Rivera *et al.*, 2010). We decided to further investigate this hypothesis through mass spectrometry and electron microscopy analyses on culture supernatants.

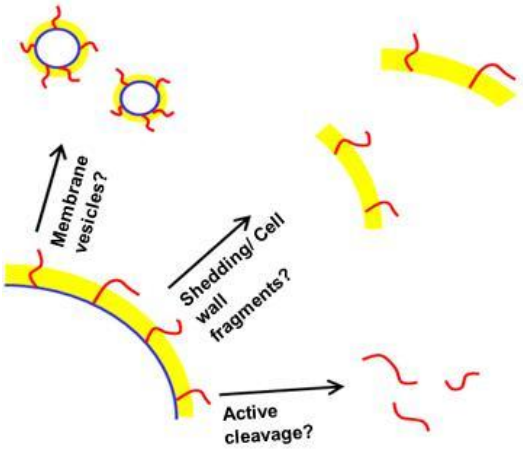


Figure 4. Scheme representing possible mechanisms by which *C. difficile* could release surface-associated proteins in the extracellular environment.

2.2. Mass spectrometry analysis of 630 supernatant ultracentrifugation fractions

We hypothesized that heavy macromolecular structures, such as cell wall fragments or membrane vesicles, should be spinned down in the pellet during ultracentrifugation of culture supernatants. To evaluate this hypothesis, *C. difficile* 630 was grown in chemically defined medium to exponential phase. After removal of all the bacterial cells, culture supernatant was ultracentrifuged to obtain two fractions: an ultracentrifugation supernatant (UC SN) and an ultracentrifugation pellet (UC P). Mass spectrometry analysis of the protein content was carried out on both fractions as well as on the total supernatant in two experiments (Table 4).

The protein content of culture supernatant before and after ultracentrifugation was comparable, suggesting that most proteins are not released in association with large structures (Table 4). In accordance with this observation, we detected few proteins in the ultracentrifugation pellet (Table 4). Among them, we could detect the S-layer proteins, which are very abundant on the bacterial surface, flagellar proteins and a CWP (Cwp24).

Family	Gene annotation (630)	Locus tag 630	exp5 TOT SN	exp5 UC SN	exp5 UC P	exp6 TOT SN	exp6 UC SN	exp6 UC P
CWPs	precursor of the S-layer proteins (SlpA)	CD630_27930	X	X	X	X	X	X
	cell wall binding protein (Cwp2)	CD630_27910					X	
	cell wall binding protein (Cwp11)	CD630_27950					X	
	hemagglutinin/adhesin (CwpV)	CD630_05140	X	X		X	X	
	cell surface protein penicillin-binding protein (Cwp20)	CD630_14690		X			X	
	pseudogene (Cwp24)	CD630_21930	X	X	X	X	X	
	cell surface protein (Cwp84)	CD630_27870		X				
	N-acetylmuramoyl-L-alanine amidase, autolysin (Cwp6)	CD630_27840	X	X		X	X	
	N-acetylmuramoyl-L-alanine amidase (Cwp16)	CD630_10350					X	
LPXTG-like	cell wall hydrolase	CD630_01830	X	X		X	X	
Extracellular hydrolases	cell-wall hydrolase	CD630_27680	X	X		X	X	
	peptidoglycan-binding/hydrolysing protein	CD630_23880				X	X	
	hydrolase	CD630_24800		X			X	
	LmbE-like deacetylase	CD630_27900		X		X	X	
Transporters	ABC transporter sugar-family extracellular solute-binding protein	CD630_08730	X	X		X	X	
	ABC transporter iron-family extracellular substrate-binding protein	CD630_16500	X			X		
	ABC transporter cystine/aminoacid-family extracellular solute-binding protein	CD630_21770		X		X	X	
	ABC transporter oligopeptide-family solute-binding protein	CD630_26720		X			X	
	ABC transporter nitrate/sulfonate/taurine extracellular solute-binding protein	CD630_23650		X			X	
Flagellar proteins	flagellar hook-associated protein FlgK (or HAP1)	CD630_02310	X	X		X	X	
	flagellar hook-associated protein 2 FliD (or HAP2)	CD630_02370	X	X	X	X	X	
	flagellin C	CD630_02390	X	X	X	X	X	X
	flagellar hook protein FlgE (Distal rod protein)	CD630_02550		X		X	X	
	flagellar basal body rod protein FlgG	CD630_02690	X	X		X	X	
	flagellar hook-associated protein FlgL (or HAP3)	CD630_02320	X	X		X	X	
	flagellar basal-body rod protein FlgC	CD630_02460		X		X	X	
Hypothetical proteins	hypothetical protein	CD630_36690					X	
	hypothetical protein	CD630_28300	X	X		X	X	
	hypothetical protein	CD630_21270		X		X	X	
	hypothetical protein	CD630_05490		X		X	X	
	hypothetical protein	CD630_11560				X		
Other predicted surface proteins	Signal peptidase I, S26A family	CD630_13310		X			X	
	SH3-domain-containing protein	CD630_11350		X				
	calcium-binding adhesion protein	CD630_27970					X	
Predicted cytoplasmic proteins	translation inhibitor endoribonuclease	CD630_25130					X	
	enolase (2-phosphoglycerate dehydratase) (2-phospho-D-glycerate hydro-lyase)	CD630_31700					X	
	elongation factor Tu	CD630_00580		X				

Table 4. Proteins detected in 630 ultracentrifugation fractions (experiments 5 and 6 from Table 3).

2.3. Electron microscopy analysis of 630 supernatant ultracentrifugation fractions

Ultracentrifugation pellets were analyzed by electron microscopy. Flagellar-like structures were identified, indicating that flagella are released as entire structures, probably because of surface remodeling (Fig. 5). These structures have a diameter of approximately 12 nm, compatible with a bacterial flagellum. Also, a small amount of vesicular-like structures were observed. These structures presented variable sizes, ranging from ~50 nm to ~200 nm of diameter (Fig. 5). Mass spectrometry analysis of ultracentrifugation pellet suggested that these structures are not involved in the massive release of surface proteins that we have observed (Table 4). Nevertheless, this is an interesting observation as membrane vesicles were never detected previously in clostridial culture media.

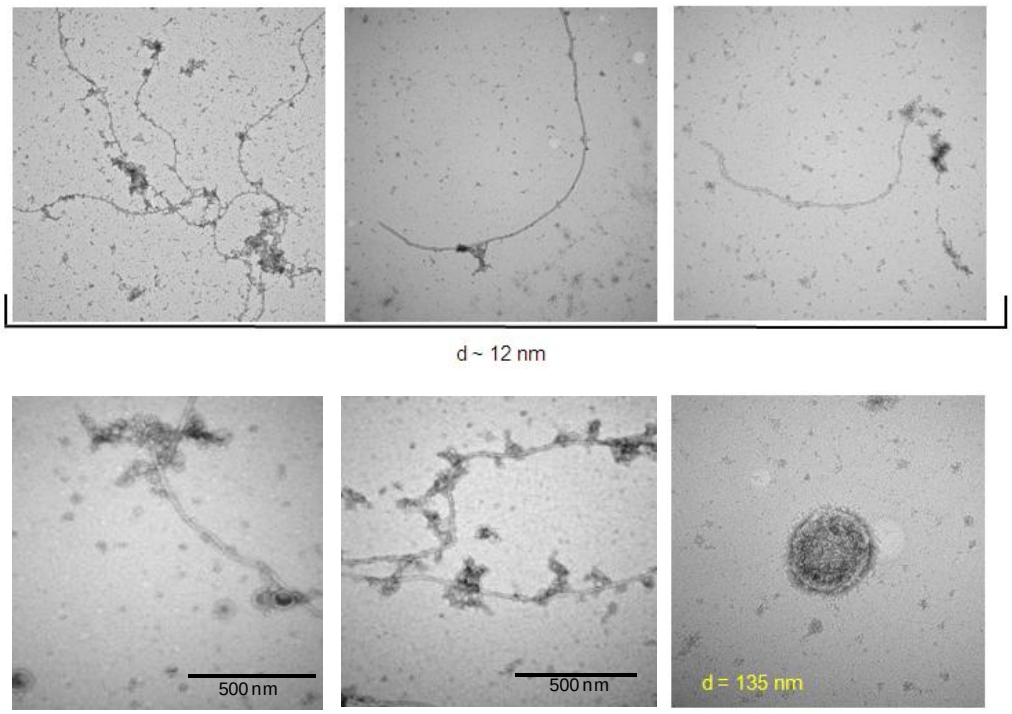


Figure 5. Electron microscopy analysis of the ultracentrifugation pellet of 630 culture supernatant. Flagellar-like structures with a diameter of ~12 nm (top) and vesicular-like structures of variable sizes (bottom) were identified.

3. Cellular localization of proteins identified by proteomics

In order to clarify the subcellular localization of the proteins found in 630 culture supernatants, we selected three proteins representing different classes and analyzed their presence in culture supernatants and in cellular fractions (Fig. 6). Antibodies against recombinant CD630_01830 (MW: 37 kDa), CD630_23650 (MW: 39kDa) and CD630_02370 (FliD, MW: 58 kDa) proteins were generated in mice. 630 cell and supernatant fractions were then analyzed by immunoblotting.

Presence in culture supernatants was confirmed for all three proteins. Particularly, CD630_01830 and CD630_23650 were found in the supernatant both before and after ultracentrifugation but were absent in the pellet, indicating that these proteins were not derived from any membranous or cell wall fragments released from the cells. On the contrary, FliD was also found in the ultracentrifugation pellet, supporting the hypothesis that entire flagellar structures may be released into the media (Fig. 5 and 6).

In accordance to the predicted association to the cell wall, CD630_01830 and CD630_23650 proteins were found in the fractions containing cell-wall associated proteins (ME and SL). On the contrary, the flagellar cap protein FliD was largely found as expected only in the protoplast cell fraction (Fig. 6).

To understand if these proteins were immunogenic in mice, recombinant proteins were probed by immunoblotting with anti-sera raised against concentrated supernatants (Fig. 7). In this analysis, in addition to the proteins described above, we included five additional recombinant proteins, namely CD630_14690 (MW 77 kDa), CD630_27950 (MW 59 kDa), CD630_05140 (MW 123 kDa), CD630_21770 (MW: 29 kDa) and CD630_36690 (MW: 21kDa). We observed bands of the predicted molecular weights for all the proteins tested, indicating that all of them were present in the supernatants and were immunogenic (Fig. 7B).

Thus our data confirm that surface-associated proteins are present in the *C. difficile* culture supernatants and at least some of these are immunogenic in mice.

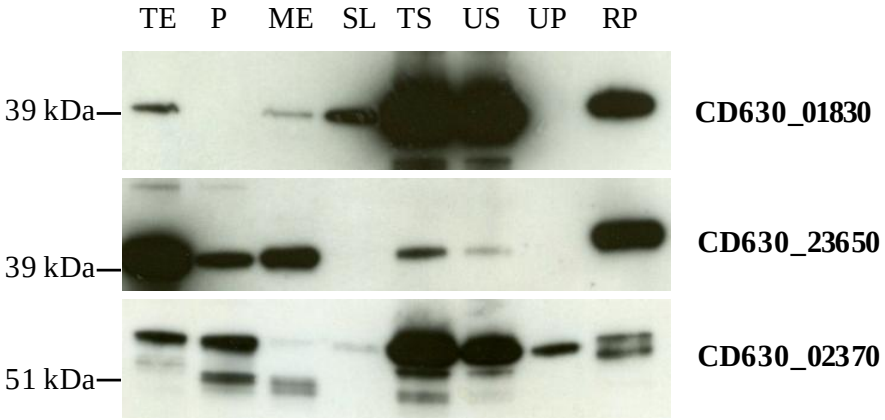


Figure 6. Cellular localization of *C. difficile* 630 proteins detected by LC-MS/MS. Immunoblotting analysis of cell and supernatant fractions prepared from exponential phase cultures with antisera generated against CD630_01830, CD630_23650 and CD630_02370 (FliD). Cell fractions: total extract (TE), protoplast (P), mutanolysin extract (ME), S-layer extract (SL). Supernatant fractions: total supernatant (TS), supernatant after ultracentrifugation (US), ultracentrifugation pellet (UP). 10 ng of each recombinant protein (RP) were loaded as control.

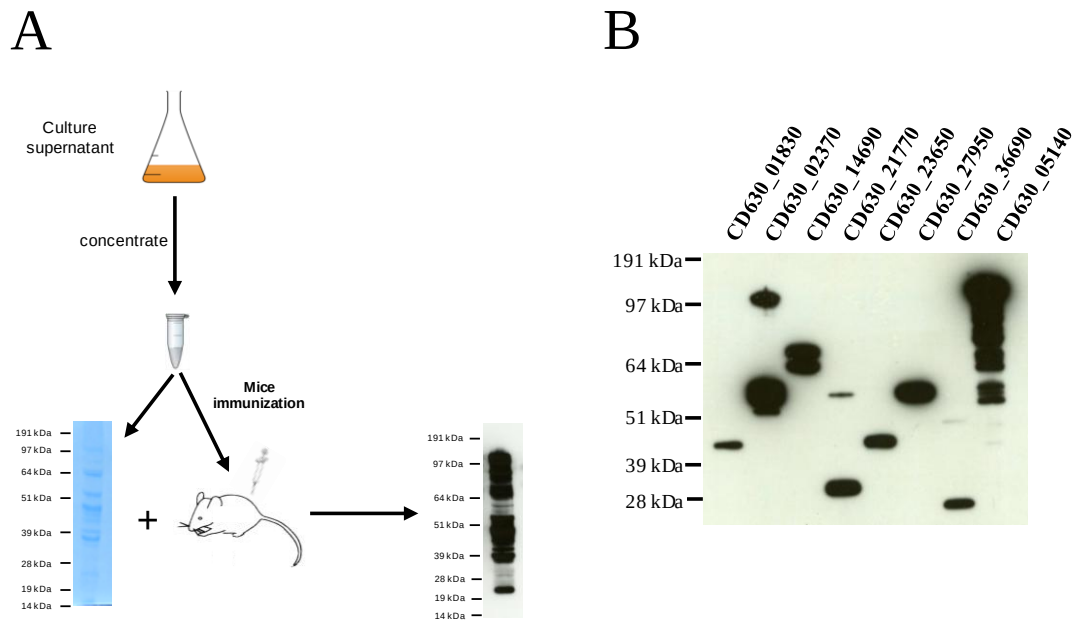


Figure 7. Immunoblotting on *C. difficile* 630 proteins detected by LC/MS/MS using an anti-supernatant antibody. (**A**) Generation of a serum raised against culture supernatants by immunization of mice with concentrated supernatants from exponentially grown bacteria. (**B**) Immunoblotting analysis of recombinant proteins (CD630_01830, CD630_02370, CD630_14690, CD630_21770, CD630_23650, CD630_27950, CD630_36690 and CD630_05140) with anti-supernatant serum.

4. Selection of proteins potentially involved in *C. difficile* pathogenesis for functional characterization

After a subset of *C. difficile* surface/released proteins was identified, we focused our attention on proteins annotated as "hypothetical", for which a putative function has not been hypothesized. 6 hypothetical proteins were identified (Table 2 and 3). 4 of them were detected in both 630 and R20291, while CD630_21270 and CD630_11560 were detected only in 630 supernatants. We took advantage of basic bioinformatics methods to identify, through analysis of the amino acid sequences of these proteins, conserved domains which could suggest a function (Table 5).

CD630_05490 was found to contain a formin homology 2 domain (pfam02181), found in proteins involved in actin polymerization. CD630_11560 contains a domain belonging to the transglutaminase-like superfamily (pfam01841); bacterial proteins containing domains of this family have been proposed to act as proteases (Makarova *et al.*, 1999). For 2 proteins, CD630_21270 and CD630_2251, no conserved domains could be identified. A blastp search of CD630_28300 (against all the proteins in the NCBI database) revealed similarity to sequences containing an Anthrax Toxin Lethal Factor domain (ATLF, pfam07737). This is the catalytic domain of *Bacillus anthracis* major toxin Lethal Factor (Tonello and Montecucco, 2009). CD630_36690 contains at the C-terminus a GERMN domain (pfam10646). GerMN is a region of approximately 100 residues that is found in a number of different species of sporulating and non-sporulating bacteria, both alone and in association with other domains (Rigden and Galperin, 2008). *Bacillus subtilis* protein GerM, in which two copies of GerMN domain are present, has been implicated in both sporulation and spore germination (Sammons *et al.*, 1987; Slynn *et al.*, 1994). Based on these analysis, we decided to focus our attention on further characterization of CD630_28300 and CD630_36690.

CD630_28300 protein was detected in both 630 and R20291 (CDR20291_2721) culture media and its amino acid sequence is conserved among the *C. difficile* genomes filed in the NCBI database with an aminoacid identity ranging from 98 to 100% (it is present in 18 out of 20 genomes, the two remaining genomes being reported as "draft assemblies"). The presence of CD630_28300 in supernatants could not be confirmed by immunoblotting. As mice serum raised against the recombinant protein was not able to recognize it, the recombinant form of this protein is probably poorly immunogenic. This hypothesis is supported also by the inability of the serum raised against the concentrated supernatant to recognize the recombinant protein (data not shown).

Blastp searches of the NCBI database using the CD630_36690 amino acid sequence as the query against each of 20 *C. difficile* strains reported in the database showed that this protein is conserved in 18 out of 20 *C. difficile* strains, with a percentage of identity ranging from 96% to 100% (the two remaining genomes being reported as "draft assemblies"). The presence of this protein in culture supernatant was confirmed by immunoblotting (Fig. 7B). At the N-terminus of CD630_36690 a signal peptide is predicted (SignalP). Also, analysis with PROSITE revealed the presence of a putative Prokaryotic lipoprotein motif (PS51257).

Hypthetical protein	**Result of *in silico* analysis**
CD630_05490	formin homology 2 domain
CD630_11560	transglutaminase-like superfamily
CD630_21270	no conserved domains
CD630_22510	no conserved domains
CD630_28300	similarity to Anthrax Lethal Factor catalytic domain
CD630_36690	GerMN domain

Table 5. Conserved domains identified by *in silico* analysis of the hypothetical proteins detected in 630 and R20291 supernatants.

5. CD630_28300 is a novel protease displaying zinc-dependent activity on the extracellular matrix protein fibronectin

5.1. CD630_28300 shares similarity to the Anthrax Lethal Factor

Alignment of the ATLF domain of CD630_28300 to the C-terminal, catalytically active, ATLF domain of Anthrax Lethal Factor (LF) of *Bacillus anthracis* showed that the two sequences share an amino acid identity of 22% and that residues mainly involved in proteolytic activity for LF and other zinc proteases are conserved, including the HEXXH motif (Fig 8).

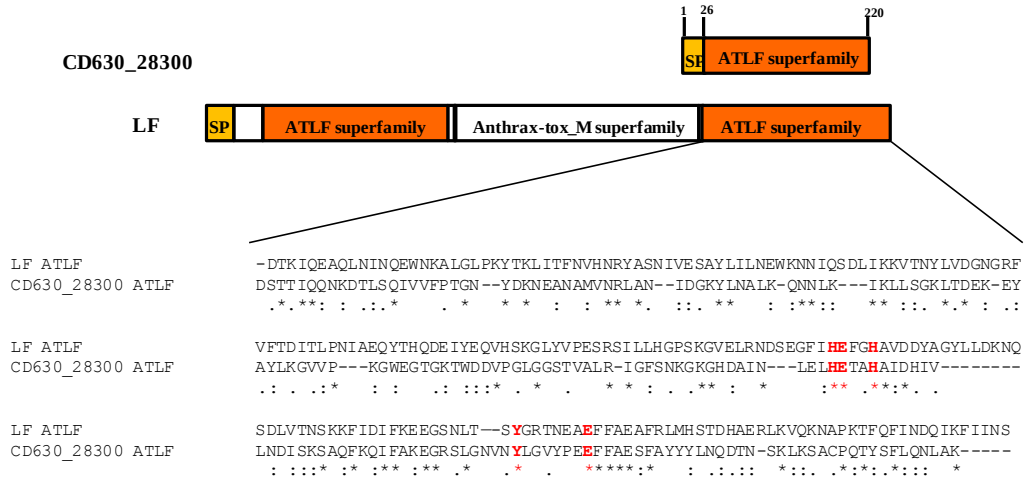


Figure 8. CD630_28300 shares sequence similarity with the C-terminal ATLF catalytic domain of Anthrax Lethal Factor (LF). Residues previously involved in the proteolytic activity in *B. anthracis* and other zinc metallopeptidases (indicated in red) are conserved. SP, signal peptide.

5.2. Preparation of recombinant CD630_28300 for activity assays

Given similarity with anthrax LF and conservation of catalytic residues, we decided to investigate proteolytic activity of CD630_28300. For this purpose, we expressed and purified a 6His-tagged recombinant protein. As the His tag is able to interact with divalent cations and could interfere with CD630_28300 metal binding ability and proteolytic activity, we decided to generate a tag-less protein using a two-step purification protocol, as described in Methods. In order to study

functionality of the protein in the absence of metals or in presence of different metals, treatment with a chelating agent (EDTA) was carried out, eventually followed by incubation with $ZnCl_2$, $CuCl_2$ or $NiCl_2$.

5.3. CD630_28300 is a novel zinc-metalloprotease

We tested catalytic activity of the recombinant protein on gelatin, a generic substrate commonly used to detect proteolytic activity, using a fluorimetric assay (Fig. 9). In this experiment, gelatin peptides are heavily labelled with both fluorescein (FITC) and a quenching agent. The enzymatic cleavage of the substrate molecule results in the separation of the quencher from the fluorochrome and consequently in an increase in fluorescence (that is proportional to the proteolytic activity). The experiment was carried out with the protein in the "apo" form or in presence of several divalent cations: Zn^{2+}, Cu^{2+} and Ni^{2+}. We could observe a weak, slow proteolytic activity in the presence of Zn^{2+}, but not in the other conditions tested (Fig. 9). The weakness of the observed activity is not surprising, as we cannot rule out that the substrate and reaction conditions used are suboptimal for this protease. However, a clear proteolytic activity could be detected, interestingly only in the presence of zinc, as we could expect based on conservation of zinc-binding residues (Fig. 8).

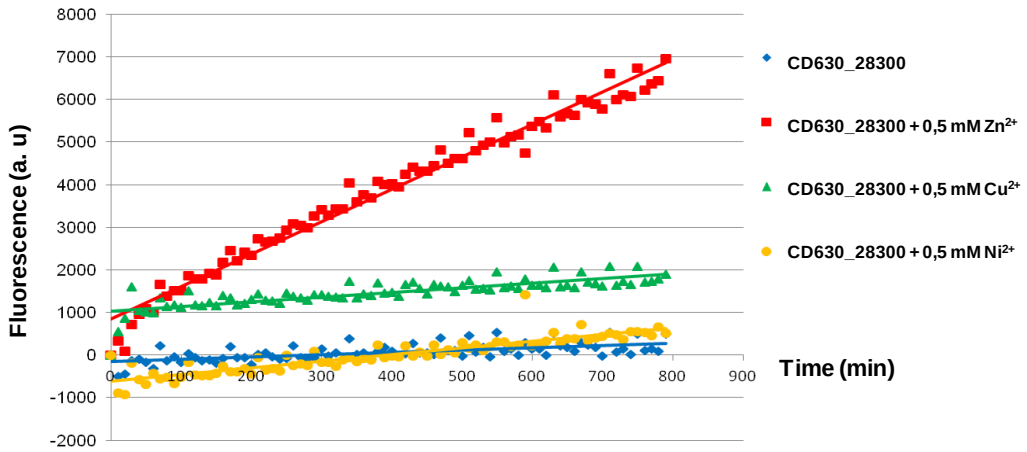


Figure 9. Proteolytic activity of CD630_28300 in a gelatinase/collagenase fluorimetric assay.

5.4. CD630_28300 displays a zinc-dependent cleavage of the extracellular matrix protein fibronectin

As other bacterial metalloproteases have been previously reported to be able to degrade extracellular matrix (ECM) proteins (Molla *et al.*, 1988; Teo *et al.*, 2003), we examined if CD630_28300 could cleave ECM components *in vitro* in the presence of Zn^{2+}. We tested collagen types I, II, III, IV, V and VI and fibronectin. While none of the collagens was cleaved (data not shown), fibronectin was cleaved in a time-dependent manner (Fig. 10A). To rule out that the observed activity was unspecific, a similar reaction was carried out in the presence of an unrelated recombinant protein purified in a similar manner (Fig. 10A). Testing CD630_28300 proteolytic activity in the absence of metals or in the presence of Cu^{2+} or Ni^{2+}, we observed the highest activity in the presence of Zn^{2+}, in accordance with the zinc-binding function previously reported for ATLF domain (Fig. 10B). These data demonstrate that CD630_28300, which we have named zinc metalloprotease 1 (Zmp1), is a novel *C. difficile* zinc-dependent protease with acivity on fibronectin.

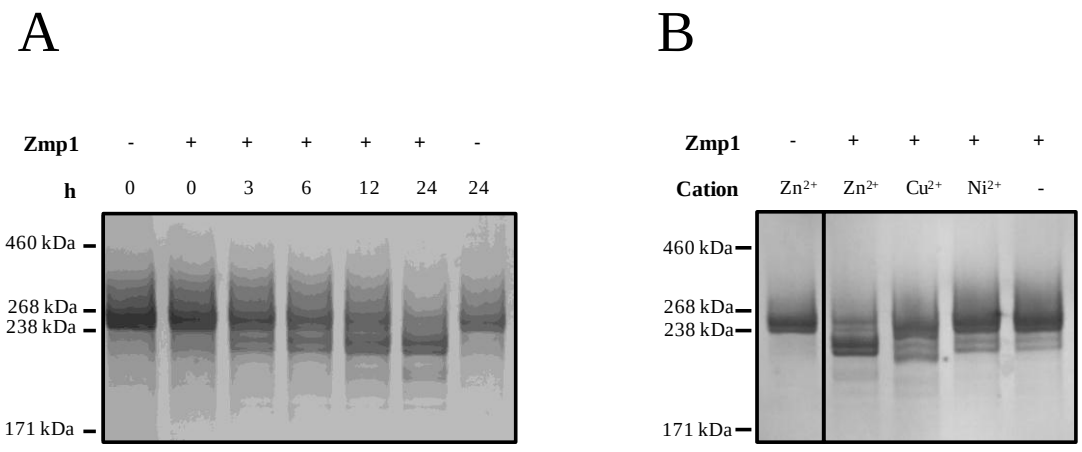


Figure 10. Zmp1 is a zinc-dependent protease with fibronectin-cleaving activity. (**A**) Time-dependent proteolytic activity of recombinant Zmp1 on fibronectin. 30 µg fibronectin from human plasma were incubated for 24 hours at 37°C with an equal amount of Zmp1 in the presence of 0.5 mM $ZnCl_2$. At 0, 3, 6, 12 and 24 hours after incubation, 1 µg of fibronectin was analyzed by SDS-PAGE followed by silver staining. Integrity of fibronectin in the absence of Zmp1 was verified after 0 and 24 h of incubation in the same conditions. (**B**) Zinc-dependent proteolytic activity of recombinant Zmp1 on fibronectin. 5 µg of fibronectin from human plasma were incubated for 24 h at 37°C with an equal amount of Zmp1 in the presence of 0.5 mM $ZnCl_2$, $NiCl_2$, $CuCl_2$ or in the absence of divalent cations. 1 µg of fibronectin was analyzed by SDS-PAGE and silver staining.

5.5. Residues E143 and H146 are important for Zmp1 catalytic activity

To further analyze the residues that are involved in Zmp1 catalytic activity, we generated mutants in the HEXXH motif, that is considered a fingerprint for zinc metalloproteases and has been characterized in anthrax LF as well as in other toxins, such as the tetanus and botulinum neurotoxins. In this motif, the imidazole rings of the two histidine residues have been identified as part of the first shell of zinc coordination, while the glutamate carboxylate binds the water molecule implicated in the hydrolytic reaction (Tonello and Montecucco, 2009).
We generated by site-directed mutagenesis two mutants of Zmp1, E143A and H146A. Stability and zinc-binding ability of these mutants were evaluated by differential scanning fluorimetry (DSF) (Fig. 11A). In the DSF assay, both mutants showed a melting temperature comparable to the wild type, indicating that mutations do not affect protein stability. The E143A mutant was able to bind zinc, as demonstrated by an increase in melting temperature observed upon addition of this metal. On the contrary, in the H146A mutant the ability to bind zinc was completely lost (Fig. 11A). Both mutant proteins showed a reduced proteolytic activity on fibronectin in the presence of zinc, indicating that residues E143 and H146 are both important for the catalytic activity (Fig. 11B).

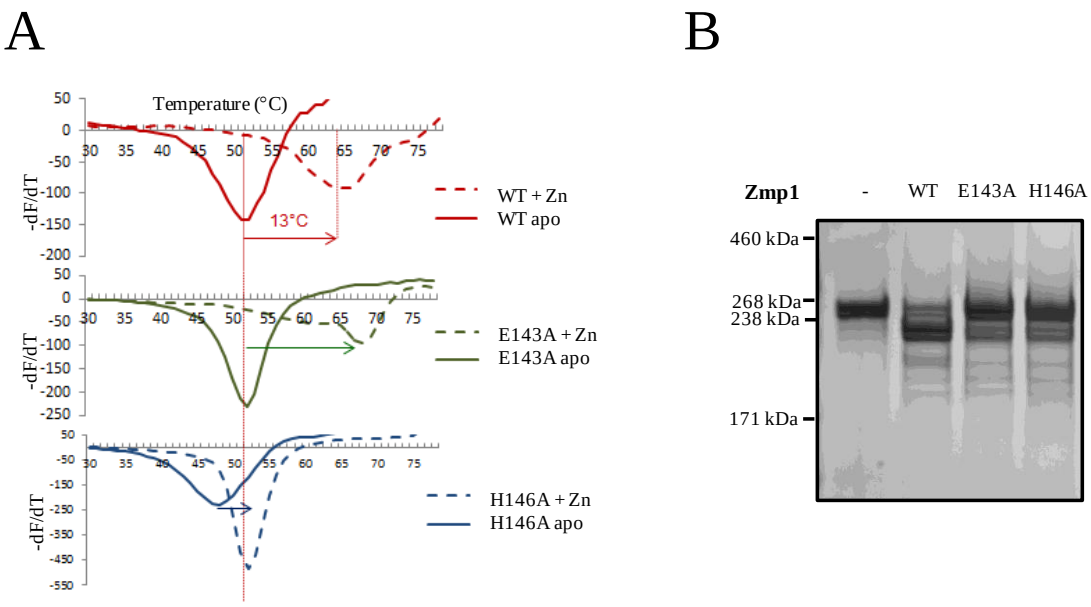


Figure 11. Decreased fibronectin-cleaving activity of E143A and H146A Zmp1 mutants. (**A**) Differential scanning fluorimetry of wild type, E143A and H146A recombinant proteins in the absence or presence of Zn^{2+}. The minimum of the derivative is the melting temperature. Shifts in melting curves for the wild type and mutant proteins are indicated with arrows. (**B**) Decreased ability of E143A and H146A Zmp1 mutants to cleave fibronectin. 5 µg of fibronectin from human plasma were incubated for 24 h at 37°C with no Zmp1 or with an equal amount of wild type (wt), E143A or H146A Zmp1 in the presence of 0.5 mM $ZnCl_2$. 1 µg of fibronectin was analyzed by SDS-PAGE and silver staining.

5.6. Zmp1 destabilizes the fibronectin network produced by human fibroblasts

We examined if the ability to cleave fibronectin *in vitro* was relevant also in a cell model resembling the ECM components produced by host cells. We used a human fibroblast cell line (IMR-90) producing a complex fiber organization mainly composed of fibronectin and collagens. Upon incubation of the cells with wild type Zmp1, we found that the discrete fibronectin network observed in the untreated cells was destabilized (Fig. 12); particularly, we observed the appearance of unstructured fibers of cellular fibronectin indicative of proteolytic activity of the protein (inset in Fig. 12). This effect is zinc-dependent since the fiber disorganization is maximized by the addition of $ZnCl_2$ in the culture, whereas a minor effect is observed in the absence of the cation in the medium

(data not shown). E143A and H146A mutations in Zmp1 abrogate the observed destabilization of the fibronectin network (Fig. 12).

These data suggest that Zmp1 is able to cleave native fibronectin and may indicate a potential role for this enzyme during *C. difficile* colonization of the gut and dissemination.

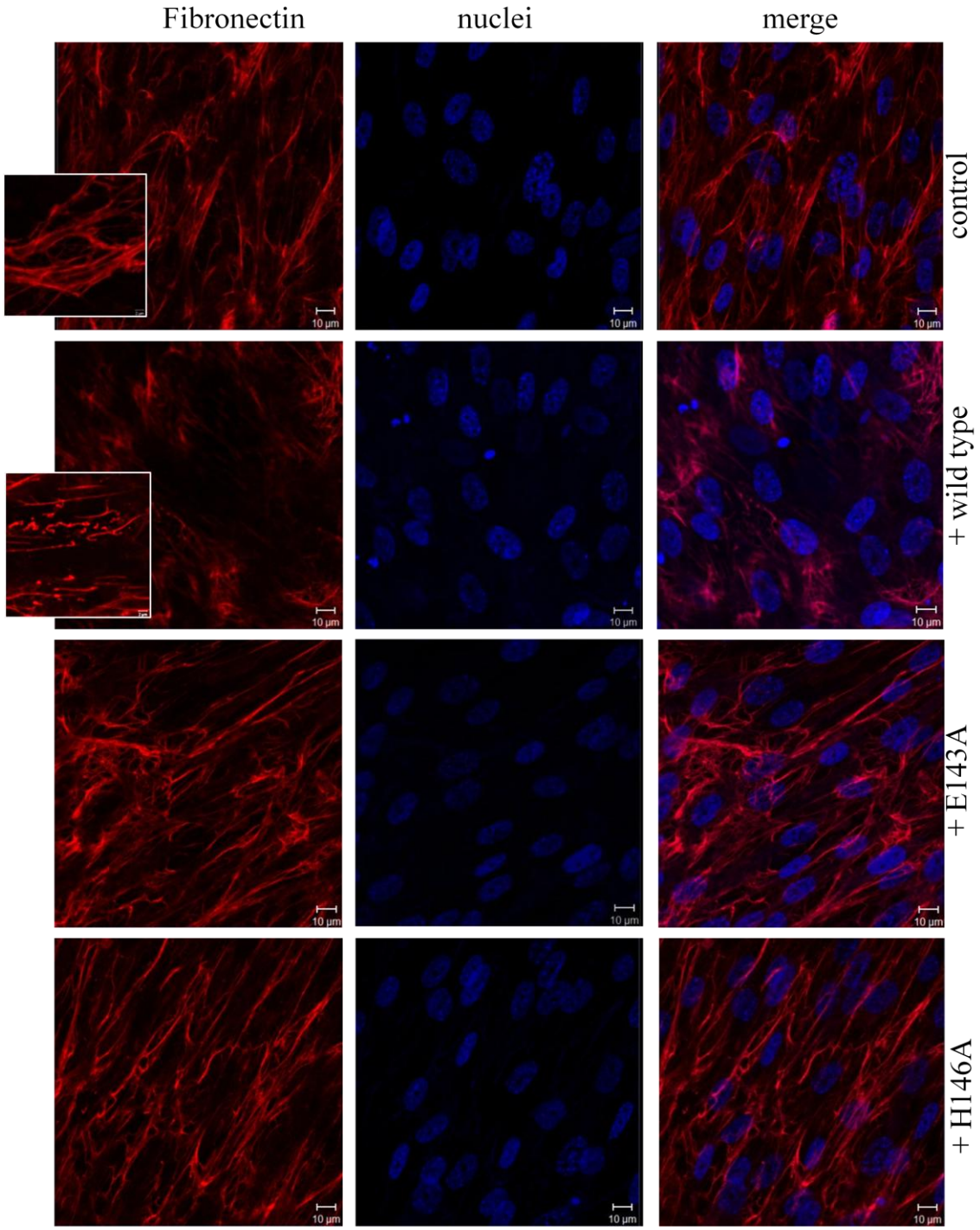


Figure 12. Proteolytic activity of Zmp1 protein on native fibronectin produced by cultured human fibroblasts. IMR-90 human fibroblasts were incubated with 250 µg/ml of wild type, E143A or H146A Zmp1 protein for 16 h as described in Methods. Fibronectin was labeled using a specific antibody followed by Alexa568-conjugated secondary antibodies (red) and nuclei were stained with DAPI (blue). Control cells were incubated with an equivalent volume of buffer for the same time period. Destabilization of fibronectin was observed upon treatment of cells with Zmp1, as also highlighted in the inset.

6. CD630_36690 is a surface protein with a potential role in *C. difficile* spores

6.1. CD630_36690 contains a C-terminal GerMN domain

CD630_36690 has a predicted amino acid sequence of 191 amino acids (Fig. 13). At the N-terminus a signal peptide is predicted (SignalP). Also, analysis of CD630_3669 with PROSITE revealed a Prokaryotic lipoprotein motif (PS51257). Searching for conserved domains on NCBI, a GERMN domain (pfam10646), a widespread domain previously implicated in sporulation and germination in *B. subtilis* GerM, was identified at the C-terminus (amino acids 74-181) (Rigden and Galperin, 2008; Slynn *et al.*, 2004) (Fig. 13). GerMN domain, which is often found associated to other domains or in two copies (as is the case of GerM protein) (Rigden and Galperin, 2008), is found alone in CD630_36690 (Fig. 13).

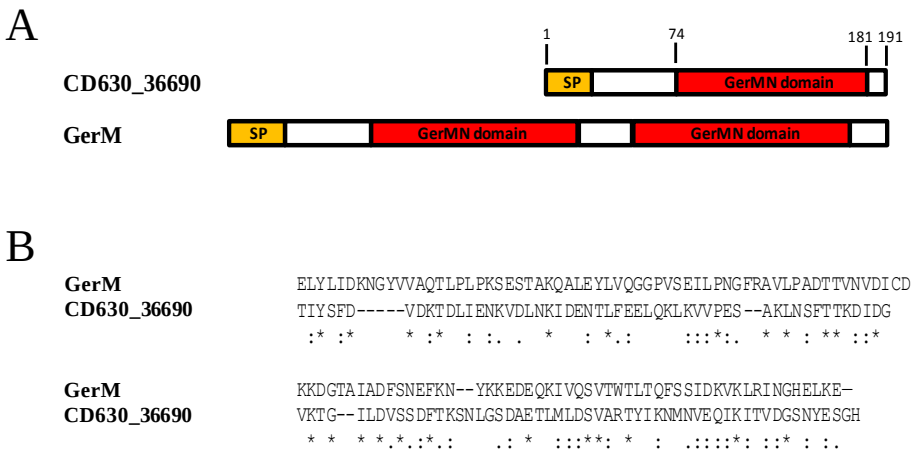


Figure 13. CD630_36690 contains a C-terminal GerMN domain. (A) Schematic representation of *C. difficile* CD630_36690 and *Bacillus subtilis* GerM proteins. **(B)** Alignment of the first GerMN domain of GerM with the GerMN domain of CD630_36690. SP, signal peptide.

6.2. Generation of a *630/ΔCD630_36690* deletion mutant

To investigate the function of CD630_36690, a deletion mutant was generated in strain 630, taking advantage of an allele exchange strategy set up by Cartman and co-workers (Cartman *et al.*, 2012).

A suitable construct was generated cloning "in frame", in the pMTL-SC7315 vector (Cartman *et al.*, 2012), the two regions flanking the gene of interest, in order to substitute the functional endogenous gene with a short sequence coding for a non-functional short peptide. The construct was transformed in *C. difficile*, where two steps of selection where carried out, as described in Methods. First, clones which had integrated the entire plasmid were selected for antibiotic resistance and checked by PCR (data not shown); then, clones which had undergone excision of the plasmid (resulting either in deletion of the gene or in recovery of the wild type genotype) were positively selected on fluorocytosin and the genotype (wild type or deletion mutant) was determined by PCR (Fig. 14).

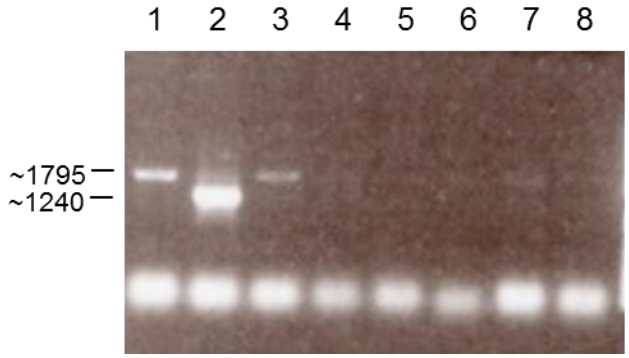


Figure 14. PCR selection of *630/ΔCD630_36690* deletion mutant. PCR was carried out on DNA extracts of fluorocytosine-positive clones. Primers external to the CD630_36690 coding sequence were used. Expected molecular weight is 1795 bp for the wild type, 1240 bp for the deletion mutant.

6.3. CD630_36690 is expressed on the surface of vegetative cells

Subcellular localization of CD630_36690 was carried out by immunoblotting and immunofluorescence analysis. Immunoblotting was performed on bacterial fractions prepared from a 630 culture collected in exponential phase, using an

antibody raised against the recombinant protein. Protoplast, cell wall and S-layer fractions of 630 and *630/ΔCD630_36690* were analyzed. CD630_36690 was mainly detected in the cell wall fraction, which contains proteins associated to the peptidoglycan as well as proteins associated to the S-layer. At least two bands of approximately 20 kDa can be observed. These bands may correspond to different maturation forms of a lipoprotein. A band of approximately 14 kDa, which can be observed only in the cell wall fraction, may correspond to a degradation product. All the bands detected are specific, as none of them can be observed in fractions of the deletion mutant prepared with the same procedure (Fig. 15A). Immunofluorescence analysis carried out on 630 and *630/ΔCD630_36690* vegetative cells collected in late exponential phase confirmed that CD630_36690 is expressed on the surface of vegetative cells (Fig. 15B).

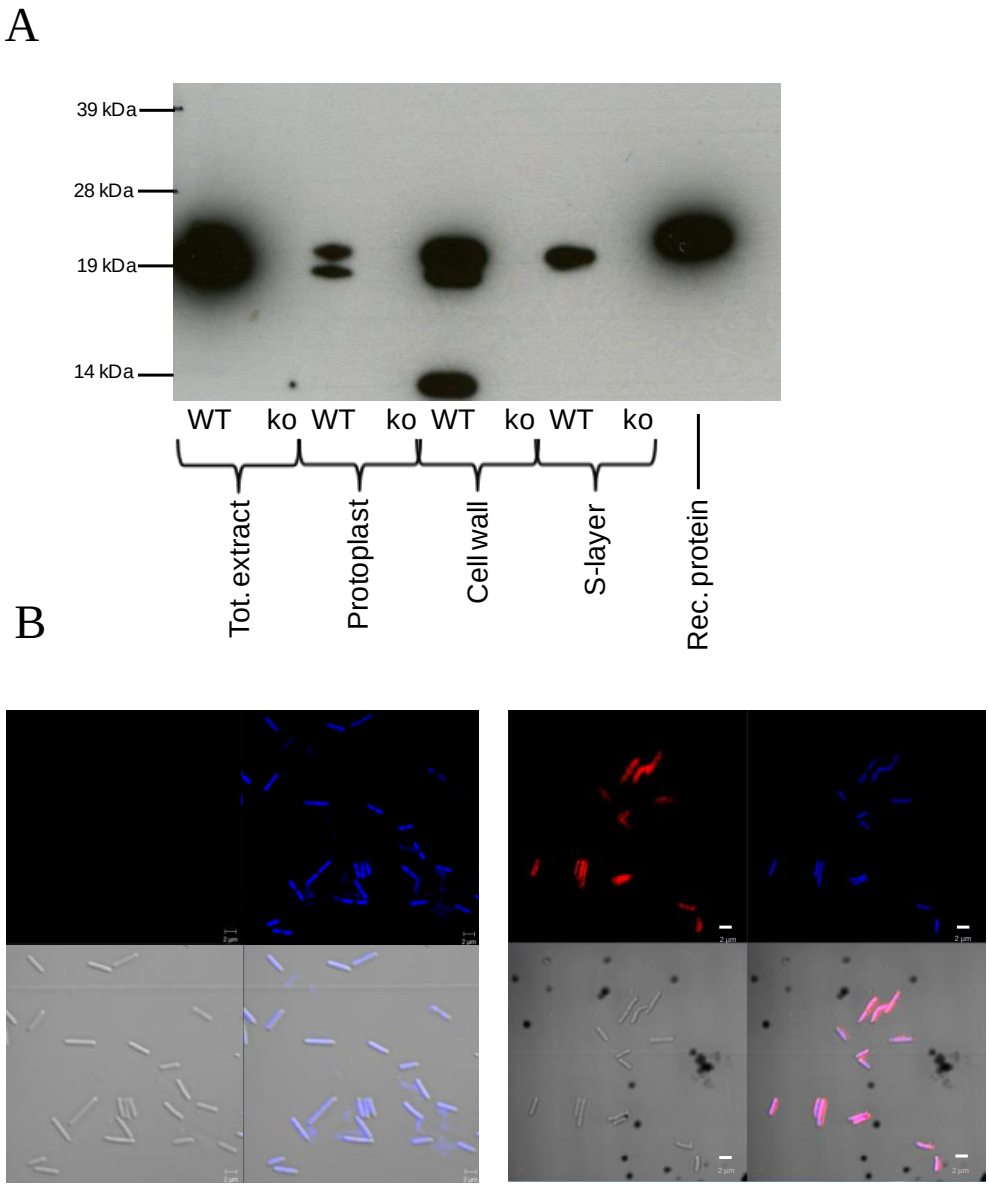


Figure 15. CD630_36690 is expressed on the surface of 630 vegetative cells. (**A**) Immunoblotting analysis of cell fractions from 630 (WT) and *630/ΔCD630_36690* (ko) using an anti-CD630_36690 specific antibody. Cell fractions: total extract, protoplast, cell wall (mutanolysin extract), S-layer extract. 10 ng of recombinant protein were loaded as a control. (**B**) Immunofluorescence analysis of *630/ΔCD630_36690* (left) and 630 (right) vegetative cells. CD630_36690 was labeled using a specific antibody followed by Alexa568-conjugated secondary antibody (red). Nuclei were stained with DAPI (blue).

6.4. CD630_36690 has a potential role in *C. difficile* spore

To characterize the deletion mutant, first we compared growth in rich medium with respect to the wild type. *630/ΔCD630_36690* did not show any growth defect up to the stationary phase (Fig. 16A). On the contrary, after reaching the

stationary phase, the mutant strain showed a decrease in optical density faster than the wild type, which could suggest a faster bacterial death (Fig. 16B).

Given the previous association of GerMN domain with sporulation and germination (Sammons *et al.,* 1987; Slynn *et al.,* 1994), we wanted to assess the involvement of CD630_36690 in these processes. For this purpose, we performed vital counts of ethanol-resistant colonies of 630 and *630/ΔCD630_36690* at different times of the growth curve (Fig. 16C and D). When *C. difficile* cultures are treated with ethanol, only spores are able to survive while vegetative cells are killed (Borriello and Honour, 1981). Plated on the germination inducer taurocholate, these spores produce colonies. The number of colonies is a result of several factors: (i) the number of spores that are present in the original culture, (ii) the ability of spores to survive ethanol treatment, (iii) the ability of spores to germinate on taurocholate and (iv) the ability of germinated spores to moltiplicate and form a visible colony. *630/*Δ*CD36690* showed a reduced ability to form ethanol resistant colonies compared to 630 at different times of growth (5 and 24 h) (Fig. 16D). The mutant did not show any defect in the number of vital vegetative cells (Fig. 16C), indicating that the observed phenotype is not due to a lower number of vegetative cells in the culture. Also, we can exclude a defect in multiplication, as mutant colonies had the same size of wild type colonies (data not shown). Therefore, it is likely that this gene has a role in sporulation (affecting the number of spores produced or their resistance to ethanol treatment) or in germination.

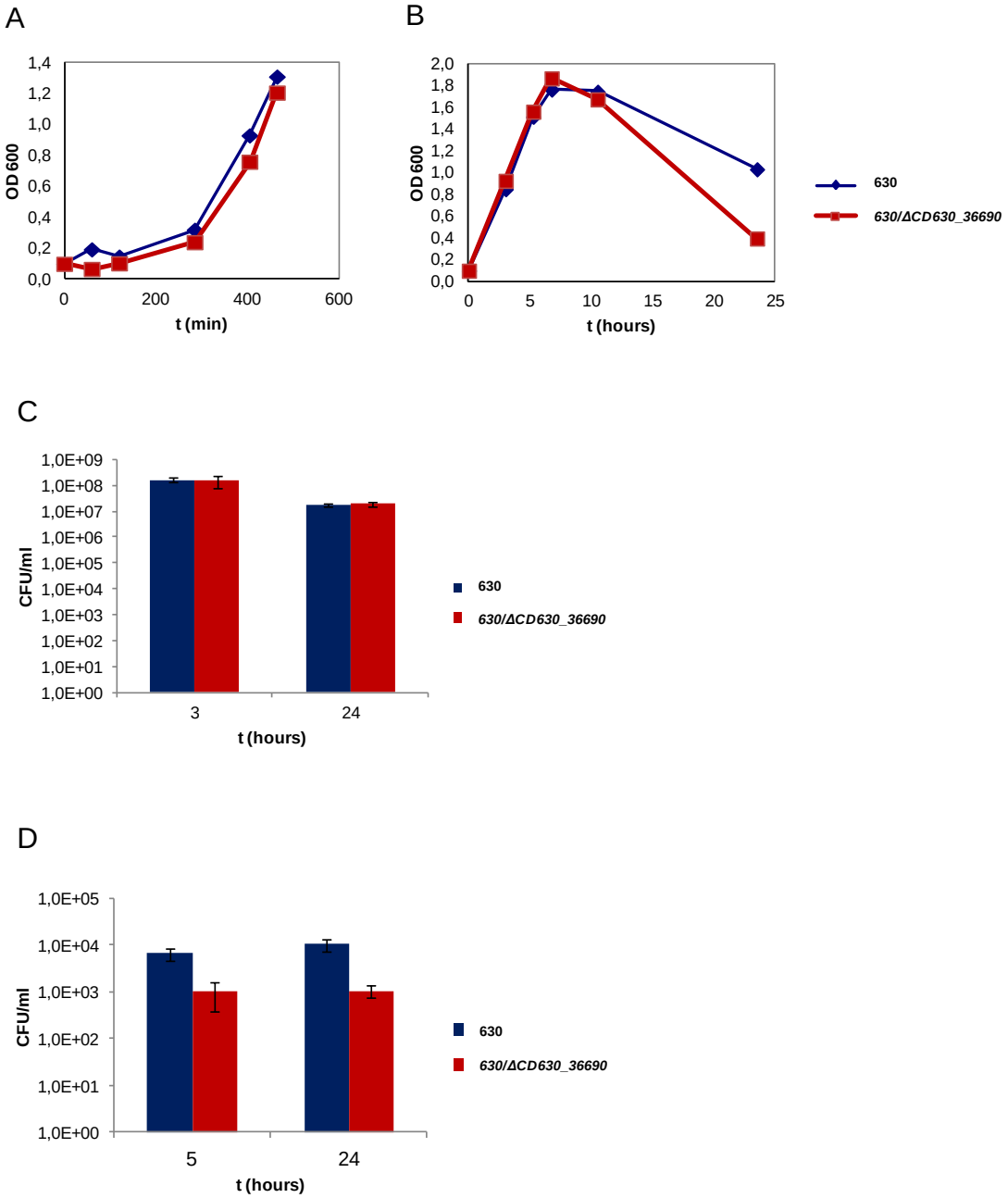


Figure 16. *630/ΔCD630_36690* has a reduced ability to form ethanol-resistant colonies with respect to 630. (**A**) and (**B**) *630/ΔCD630_36690* growth is comparable to 630 up to the exponential phase (A), but undergoes a faster decrease in optical density after reaching the stationary phase (B). (**C**) Vital counts of 630 and *630/ΔCD630_36690* vegetative cells at 3 hours (mid exponential phase) and 24 hours (death phase). Diluted cultures were plated on taurocholate. (**D**) Vital counts of ethanol-resistant colonies at 5 hours and 24 hours (earlier points were not evaluated because of the low number of spores produced). Following ethanol treatment, cultures were diluted and plated on taurocholate. All experiments were performed in triplicate.

DISCUSSION

Although extracellular proteins are known to mediate events involved in the gut pathogenesis during *C. difficile* infections, not many cell surface or secreted proteins of the nosocomial pathogen *C. difficile* have been identified or functionally characterized.

Previous studies tried to globally characterize extracellular protein profiles of this pathogen, but a certain percentage of bacterial lysis, leading to contamination of extracellular proteins preparations with cytoplasmic proteins, did not allow assignment of the identified proteins to their correct subcellular compartment (Boetzkes *et al.*, 2012; Wright *et al.*, 2005; Mukherjee *et al.*, 2002). Combining a protocol that minimizes contamination with cytoplasmic proteins and sensitive proteomic techniques, we were able to identify a subset of extracellular *C. difficile* proteins from two clinically relevant strains.

Interestingly, we found that the culture supernatant comprises largely of surface-associated proteins. This was recently reported also by Boetzkes *et al.*, who could identify the presence of several putative surface-associated proteins in culture filtrates, however along with the presence of a large number of cytoplasmic proteins (Boetzkes *et al.*, 2012). The presence of cytoplasmic proteins, which is a common indicator of lysis, is found in several proteomic studies examining the exoproteomes and complicates the interpretation of the cellular localization of proteins (Pacheco *et al.*, 2011; Sengupta *et al.*, 2010). Our work suggests that the proteins we detected in culture supernatants are indeed extracellular, as the preparations we used almost completely lacked proteins of predicted cytoplasmic origin. Among the proteins we detected in supernatants, the Cwp84 cysteine protease and SlpA, the precursor of the two major S-layer proteins, have been well studied and previously reported to be present in culture supernatants (de la Riva *et al.*, 2011). The largest known CWP, CwpV, which promotes *C. difficile*

aggregation, was also detected (Reynolds *et al.,* 2011). Several proteins that we found, such as Cwp84, SlpA, FliC, FliD, Cwp2, Cwp18, Cwp66, CD630_26720 and CD630_08730, have been shown to be immunoreactive to human sera (Pechine *et al.,* 2005; Wright *et al.,* 2008). In addition to known proteins, many other uncharacterized CWPs, hydrolases, transporters and hypothetical proteins are reported in this study. Further characterization of these proteins could provide insight into both bacterial physiology and pathogenesis.

Based on our observations, most surface-proteins identified appear to be released as single proteins and not in association to macromolecular structures (i.e. membrane vesicles or cell wall fragments). The only structure that seems to be released as a whole unit is the flagellum. Surface proteins have been reported from culture supernatants of other Gram-positive bacteria (Desvaux *et al.,* 2010; Ravipaty and Reilly, 2010). The cleavage of surface proteins by proteases has been previously proposed as a mechanism by which proteins are released from the surface in bacteria such as *Streptococcus pyogenes* (Nelson *et al.,* 2011). The mechanisms controlling the release of surface associated proteins in *C. difficile* are at present unclear. *C. difficile* encodes at least two cysteine proteases, Cwp84 and Cwp13, which are involved in surface maturation (Kirby *et al.,* 2009; Dang *et al.,* 2010; de la Riva *et al.,* 2011). A possibility is that these, other reported or yet unidentified proteases are involved in cleavage of some surface proteins of *C. difficile*. Indeed, surface proteins may be released or shed into the extracellular environment, perhaps as a result of abundant expression on the surface. Presence of surface-associated proteins in the extracellular milieu may suggest that *in vivo,* during infection, such proteins are not just associated to the bacterial surface but may also act distally.

Preliminary analyses on the mechanisms involved in release of surface proteins led us to the observation of membrane vesicles in culture supernatants. Although we did not go into the characterization of these structures, this is an interesting observation, as such structures were never observed before in clostridial

supernatants. A variety of Gram-negative pathogenic or environmental bacteria secrete outer membrane vesicles (OMVs), which are produced during normal bacterial growth (Ellis & Kuehn, 2010). These are spherical nanovesicles that have been demonstrated to contain virulence factors in several pathogenic species, like *Escherichia coli* and *Pseudomonas aeruginosa* (Wai *et al.*, 2003; Kesty & Kuehn, 2004; Bomberger *et al.*, 2009), and their contribution to bacterial pathogenesis has been characterized. In contrast, little is known about such processes in Gram-positive bacteria. Membrane vesicles production has been reported in *Staphilococcus aureus* (Gurung *et al.*, 2011; Lee *et al.*, 2009; Hong *et al.*, 2011), *Bacillus subtilis* (Dorward and Garon, 1990), *Bacillus cereus* (Dorward and Garon, 1990) and *Bacillus anthracis* (Rivera *et al.*, 2010). Interestingly, vesicles of *B. anthracis* have been involved in secretion and targeting of the anthrax major toxins (Rivera *et al.*, 2010). It is still unclear the mechanism by which *C. difficile* toxin A and toxin B are secreted. It has been proposed that these huge toxins are released through the product of the tcdE gene, present within the pathogenicity locus, which presents sequence and structure similarities to bacteriophage-encoded holins. However, it has recently been demonstrated that TcdE does not exhibit pore-forming function in *C. difficile* (Olling *et al.*, 2012). It would be interesting to study if *C. difficile* vesicles contain toxins and have a role in their targeting, as it has been shown for *B. anthracis*.

The global analysis of *C. difficile* supernatants resulted in the identification of six unknown hypothetical proteins. Based on similarities to known domains or proteins, we found some proteins that may be involved in clostridial pathogenesis. One of such proteins is the zinc-dependent metalloprotease Zmp1. We demonstrated that Zmp1 can degrade fibronectin not only *in vitro*, but also on cultured human fibroblasts that produce a dense fibronectin network. By selectively mutating residues that are in the HEXXH motif, a conserved motif typical to zinc metalloproteases and crucial for proteolytic activity (Klimpel *et*

al., 1994), we showed that specific residues E143 and H146 are crucial for the catalytic activity of Zmp1. DSF analysis showed that H146 is required for binding to Zn^{2+}, while both the residues are key for activity on purified fibronectin *in vitro*. This is in accordance with the role previously attributed to these residues in other zinc metalloproteases (Klimpel *et al.,* 1994; Teo *et al.,* 2003). Our *in vitro* studies showed that Zmp1 is active on fibronectin, but not on any of the collagen types that were tested. Moreover, a weak activity on gelatin could be detected by FRET. Thus, although Zmp1 does not demonstrate a strong fibronectin cleaving activity in our *in vitro* conditions, it is evident that this activity is highly specific. It is likely that this enzyme has optimal activity on other biological substrates. Further studies with a wider collection of known biological substrates could shed light on alternate substrates of this novel metalloprotease.

Proteolytic enzymes are frequently involved in the bacterial colonization process, contributing to nutrient acquisition, degradation of host proteins or processing of bacterial proteins involved in pathogenesis (Maeda, 1996). Previous studies showed that *C. difficile* displays a proteolytic activity correlated with strain virulence in the hamster model (Poilane *et al.,* 1998; Seddon *et al.,* 1990; Seddon and Borriello, 1992). However, few proteolytic factors of *C. difficile* have been characterized so far and the role of hydrolytic enzymes in *C. difficile* pathogenesis is still unclear. Fibronectin degrading enzymes have been previously implicated in bacterial virulence (Finkelstein *et al.,* 1983; Guillemet *et al.,* 2010). These enzymes have been shown to affect the ability to colonize in the gut, and hence to affect disease pathogenesis (Finkelstein *et al.,* 1983). During pathogenesis of *C. difficile*, Zmp1 may aid the bacterial colonization processes in the gut by degrading ECM proteins associated with the gut epithelial cells. Indeed, this enzyme may have additional biological substrates that assist in gut penetration or nutrient acquisition.

Analysis of proteins of unknown function contained in culture supernatants revealed also the presence of a novel protein containing a GerMN domain, named CD630_36690 in strain 630. Genetic studies on the *Bacillus subtilis* GerM protein, which contains two copies of GerMN domain, have previously involved this domain in both sporulation and spore germination, suggesting an important role in cell development. However, the role and mechanism of action of GerM in these processes has not been characterized (Slynn *et al.*, 1994; Sammons *et al.*, 1987). Recently, a bioinformatic analysis was carried out to study the presence of the GerMN domain in several bacterial phyla (Rigden and Galperin, 2008). These searches revealed that the organization of this domain and its combination with other conserved domains is much variable and that the GerMN domain is conserved among both sporigenous and asporigenous bacterial species. These analyses led the authors to hypothesize that GerMN is a widespread cell development domain, whose function might involve binding to the peptidoglycan. We demonstrated that this protein is expressed on the surface of vegetative cells. Moreover, our preliminary genetic study, aimed at analyzing the involvement of CD630_36690 in *C. difficile* sporulation and germination, suggested that this factor is involved in one or both of these processes. Therefore we can hypothesize two scenarios: i) the protein is a sporulation and/or germination factor, directly involved in sporulation and/or germination; ii) the protein has a general role in cell development (for example in synthesis or degradation of peptidoglycan) and deletion of the encoding gene indirectly leads to a defect in sporulation or germination. Further studies are necessary to understand in which biological processes this protein is involved and if it has a role in pathogenesis.

Thus, employing proteomics, we identified new extracellular factors that may be important for the pathogenesis of *C. difficile*. Further examination of the mechanisms involved in release of surface-associated proteins into the

extracellular milieu may shed light on how the clostridial cell surface may be modulated and how these proteins may interact with the host cells.

APPENDIX 1 – Growth media

BHIS

BHI	37 g/l
Yeast extract	5 g/l
L-cysteine	1 g/l
ddH_2O	

15 g/l agar are added to have a solid medium.

The medium is sterilized in autoclave and stored at 4°C.

TYM

Tryptone peptone (Difco)	24 g/l
Yeast extract	12 g/l
Mannitol	10 g/l
Glycerol	1 g/l
ddH_2O	

The pH is adjusted to 6.8.

The medium is sterilized in autoclave and stored at 4°C.

CDMM-*Clostridium difficile* Minimal medium

Aminoacids:

L-cysteine	0.5 g/l
L-isoleucine	0.1 g/l
L-leucine	1.0 g/l
L-proline	0.8 g/l
L-tryptophan	0.1 g/l
L-valine	0.1 g/l
L-arginine	0.1 g/l
Glycine	0.1 g/l
L-histidine	0.1 g/l
L-methionine	0.1 g/l
L-threonine	0.1 g/l

Salts:

KH_2PO_4	0.3 g/l
Na_2HPO_4	1.5 g/l
NaCl	0.9 g/l
$CaCl_2 \cdot 2H_2O$	26 mg/l
$MgCl_2 \cdot 6H_2O$	20 mg/l
$MnCl_2 \cdot 4H_2O$	10 mg/l
$(NH_4)_2SO_4$	0.44 g/l
$FeSO_4 \cdot 7H_2O$	4 mg/l
$NaHCO_3$	5 g/l

Vitamins:

Biotin	10 µg/l
Calcium pantothenate	1 mg/l
Pyridoxine hydrochloride	100 µg/l

Powders are dissolved in ddH_2O.

The pH is adjusted to 7.4.

The medium is filter sterilized and stored at 4°C.

Minimal medium for selection of double cross-over clones (CDM)

To make 100 ml Minimal medium:

Sterile anaerobic ddH_2O	61 ml
Amino acid solution (5X)	20 ml
Salt solution (10X)	10 ml
20% glucose solution	5 ml
Trace salt solution (50X)	2 ml
$FeSO_4 \cdot 7H_2O$ solution (100X)	1 ml
Vitamin solution (100X)	1 ml

Amino acid solution (5X):

Dissolve the following quantities in 200 ml of ddH_2O:

Cas-amino acids	10.0 g
L-Tryptophan	0.5 g
L-Cysteine	0.5 g

Salt solution (10X):

Dissolve the following quantities in 200 ml of ddH_2O:

KH_2PO_4	1.8 g
NaCl	1.8 g
Na_2HPO_4	10.0 g
$NaHCO_3$	10.0 g

Glucose solution:

20% w/v glucose

Trace salt solution (50X):

Dissolve the following quantities in 200 ml of ddH_2O:

$CaCl_2 \cdot 2H_2O$	260 mg
$MnCl_2 \cdot 4H_2O$	100 mg
$MgCl_2 \cdot 6H_2O$	200 mg
$(NH_4)2SO_4$	400 mg

$FeSO_4 \cdot 7H_2O$ solution (100X):

Dissolve 20 mg $FeSO_4 \cdot 7H_2O$ in 50 ml anaerobic water

Vitamin solution (100X):

Dissolve each of the following vitamins in 200 ml of water:

Ca-D-panthotenate	20 mg
Pyridoxine	20 mg
d-biotin	20 mg

The medium is filter sterilized and stored at 4°C.

LB

Tryptone peptone	10 g/l
Yeast extract	5 g/l
NaCl	10 g/l
dH_2O	

The pH is adjusted to 7.6, then the medium is sterilized in autoclave and stored at 4°C.

15 g/l agar are added to have a solid medium.

APPENDIX 2 - Oligonucleotides

E143Af	GAATTACATGCAACAGCACATGCAATAGACC
E143Ar	GTGCTGTTGCATGTAATTCTAAATTTATTGC
H146Af	GAAACAGCAGCTGCAATAGACCATATAGTATTAAATGAT
H146Ar	TATTGCAGCTGCTGTTTCATGTAATTCTAAATTTATTGC
U3669f	TTTTTTGTTTAAACCGCTCTGTAGTATTAAATTATCAGTTATC
U3669r	GGGCTAGAATTCCATAAAAATTTCTCCTCTCTTAAAAACTATTC
D3669f	GGGCTAGAATTCGGAGATTACTTAAAGTAATTTCTTGAAACT
D3669r	TTTTTTGTTTAAACCCAGGTATAGCAGGTCTTAACC
int3669f	GCTTGATTCAGTTGCTCGTACTTAC
int3669r	GAGCATTAGTACAACCAACTGCGAAT
ext3669f	CATCTGCCATAGTTCTCCATTAAATCG
ext3669r	GAAGGTGGAACAGGTAGTAAATCAG
SC7-F	GACGGATTTCACATTTGCCGTTTTGTAAACGAATTGCAGG
SC7-R	AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG
SC7-F1	GCCGTTTTGTAAACGAATTGCAGG
SC7-R1	CTACGGGGTCTGACGCTC

BIBLIOGRAPHY

Aas J., Gessert C.E., Bakken J.S. (2003). Recurrent *Clostridium difficile* colitis: case series involving 18 patients treated with donor stool administered via a nasogastric tube. *Clin Infect Dis.* ***36****:580-5.*

al Saif N., Brazier J.S. (1996). The distribution of *Clostridium difficile* in the environment of South Wales. *J Med Microbiol.* ***45****:133-7.*

Albesa-Jové D., Bertrand T., Carpenter E.P., Swain G.V., Lim J., Zhang J., Haire L.F., Vasisht N., Braun V., Lange A., von Eichel-Streiber C., Svergun D.I., Fairweather N.F., Brown K.A. (2010). Four distinct structural domains in *Clostridium difficile* toxin B. *J Mol Biol.* ***396****:1260-70.*

Allsopp L.P., Totsika M., Tree J.J., Ulett G.C., Mabbett A.N., Wells T.J., Kobe B., Beatson S.A., Schembri M.A. (2010). UpaH is a newly identified autotransporter protein that contributes to biofilm formation and bladder colonization by uropathogenic *Escherichia coli* CFT073. *Infect Immun* ***78:****1659-69.*

Aronsson B., Möllby R., Nord C.E. (1985). Antimicrobial agents and *Clostridium difficile* in acute enteric disease: epidemiological data from Sweden, 1980-1982. *J Infect Dis.****151****:476-81*.

Aubry A., Hussack G., Chen W., Kuo Lee R., Twine S.M., Fulton K.M., Foote S., Carrillo C.D., Tanha J., Logan S.M. (2012). Modulation of toxin production by the flagellar regulon in *Clostridium difficile*. *Infect Immun.* ***80****:3521-32.*

Barketi-Klai A., Hoys S., Lambert-Bordes S., Collignon A., Kansau I. (2011). Role of fibronectin-binding protein A in *Clostridium difficile* intestinal colonization. *J Med Microbiol* ***60:****1155-61*.

Bartlett J.G. (1994). *Clostridium difficile*: history of its role as an enteric pathogen and the current state of knowledge about the organism. *Clin Infect Dis. 18 Suppl 4:S265-72.*

Beloin C., Roux A., Ghigo J. M. (2008). *Escherichia coli* biofilms. *Curr Top Microbiol Immunol* ***322:****249-89.*

Bensi G., Mora M., Tuscano G., Biagini M., Chiarot E., Bombaci M., Capo S., Falugi F., Manetti A. G., Donato P., Swennen E., Gallotta M., Garibaldi M., Pinto V., Chiappini N., Musser J. M., Janulczyk R., Mariani M., Scarselli M., Telford J. L., Grifantini R., Norais N., Margarit I., Grandi G. (2012). Multi high-throughput approach for highly selective identification of vaccine candidates: the Group A Streptococcus case. *Mol Cell Proteomics **11:**M111 015693.*

Berlanda Scorza F., Doro F., Rodriguez-Ortega M.J., Stella M., Liberatori S., Taddei A.R., Serino L., Gomes Moriel D., Nesta B., Fontana M.R., Spagnuolo A., Pizza M., Norais N., Grandi G. (2008). Proteomics characterization of outer membrane vesicles from the extraintestinal pathogenic Escherichia coli DeltatolR IHE3034 mutant. *Mol Cell Proteomics **7**: 473-85.*

Best E.L., Fawley W.N., Parnell P., Wilcox M.H. (2010). The potential for airborne dispersal of *Clostridium difficile* from symptomatic patients. *Clin Infect Dis. **50**:1450-7.*

Boetzkes A., Felkel K. W., Zeiser J., Jochim N., Just I., Pich A. (2012). Secretome analysis of *Clostridium difficile* strains. *Arch Microbiol **194:**675-87.*

Bomberger J.M., Maceachran D.P., Coutermarsh B.A., Ye S., O'Toole G.A., Stanton B.A. (2009). Long-distance delivery of bacterial virulence factors by *Pseudomonas aeruginosa* outer membrane vesicles. *PLoS Pathog. **5**:e1000382.*

Borriello S.P., Honour P. (1981). Simplified procedure for the routine isolation of *Clostridium difficile* from faeces. *J Clin Pathol. **34**:1124-7.*

Borriello S.P., Welch A.R., Barclay F.E., Davies H.A. (1988). Mucosal association by *Clostridium difficile* in the hamster gastrointestinal tract. *J Med Microbiol. **25**:191-6.*

Calabi E., Calabi F., Phillips A.D., Fairweather N.F. (2002a). Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. *Infect Immun **70:**5770-8.*

Calabi E., Fairweather N. (2002b). Patterns of sequence conservation in the S-Layer proteins and related sequences in *Clostridium difficile*. *J Bacteriol. **184**:3886-97.*

Calabi E., Ward S., Wren B., Paxton T., Panico M., Morris H., Dell A., Dougan G., Fairweather N. (2001). Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Mol Microbiol **40:**1187-99.*

Carroll K.C., Bartlett J.G. (2011). Biology of *Clostridium difficile*: implications for epidemiology and diagnosis. *Annu Rev Microbiol.* ***65****:501-21.*

Carter G.P., Rood J.I., Lyras D. (2012). The role of toxin A and toxin B in the virulence of *Clostridium difficile*. *Trends Microbiol.* ***20****:21-9.*

Cartman S.T., Kelly M.L., Heeg D., Heap J.T., Minton N.P. (2012). Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the tcdC genotype and toxin production. *Appl Environ Microbiol.* ***78****:4683-90.*

Cerquetti M., Molinari A., Sebastianelli A., Diociaiuti M., Petruzzelli R., Capo C., Mastrantonio P. (2000). Characterization of surface layer proteins from different *Clostridium difficile* clinical isolates. *Microb Pathog.* ***28****:363-72.*

Cerquetti M., Serafino A., Sebastianelli A., Mastrantonio P. (2002). Binding of *Clostridium difficile* to Caco-2 epithelial cell line and to extracellular matrix proteins. *FEMS Immunol Med Microbiol.* ***32****:211-8.*

ChapetónMontes D., Candela T., Collignon A., Janoir C. (2011). Localization of the *Clostridium difficile* cysteine protease Cwp84 and insights into its maturation process. *J Bacteriol.* ***193****:5314-21.*

Cohen S.H., Gerding D.N., Johnson S., Kelly C.P., Loo V.G., McDonald L.C., Pepin J., Wilcox M.H. (2010). Clinical practice guidelines for *Clostridium difficile* infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). *Society for Healthcare Epidemiology of America; Infectious Diseases Society of America. Infect Control Hosp Epidemiol.* ***31****:431-55.*

Cohen S.H., Tang Y.J., Silva J. Jr. (2000). Analysis of the pathogenicity locus in *Clostridium difficile* strains. *J Infect Dis.* ***181****:659-63.*

Curry S.R., Marsh J.W., Muto C.A., O'Leary M.M., Pasculle A.W., Harrison L.H. (2007). tcdC genotypes associated with severe TcdC truncation in an epidemic clone and other strains of *Clostridium difficile*. *J Clin Microbiol.****45****:215-21.*

Dang T.H., de la Riva L., Fagan R.P., Storck E.M., Heal W.P., Janoir C., Fairweather N.F., Tate E.W. (2010). Chemical probes of surface layer biogenesis in *Clostridium difficile*. *ACS Chem Biol.* ***19****:279-85.*

Dapa T., Leuzzi R., Ng Y.K., Baban S.T., Adamo R., Kuehne S.A., Scarselli M., Minton N.P., Serruto D., Unnikrishnan M. (2012). Multiple factors modulate biofilm formation by the anaerobic pathogen *Clostridium difficile*. *J Bacteriol.* ***195****:545-55.*

Davies H.A., Borriello S.P. (1990). Detection of capsule in strains of *Clostridium difficile* of varying virulence and toxigenicity. *Microb Pathog.****9****:141-6.*

Dawson L.F., Valiente E., Faulds-Pain A., Donahue E.H., Wren B.W. (2012). Characterisation of *Clostridium difficile* biofilm formation, a role for Spo0A. *PLoS One* ***7****:e50527.*

Dawson L.F., Valiente E., Wren B.W. (2009). *Clostridium difficile* - a continually evolving and problematic pathogen. *Infect Genet Evol.* ***9****:1410-7.*

de la Riva L., Willing S.E., Tate E.W., Fairweather N.F. (2011). Roles of cysteine proteases Cwp84 and Cwp13 in biogenesis of the cell wall of *Clostridium difficile*. *J Bacteriol* ***193:****3276-85.*

Desvaux M., Dumas E., Chafsey I., Chambon C., Hebraud M. (2010). Comprehensive appraisal of the extracellular proteins from a monoderm bacterium: theoretical and empirical exoproteomes of *Listeria monocytogenes* EGD-e by secretomics. *J Proteome Res* ***9:****5076-92.*

Dhalluin A., Bourgeois I., Pestel-Caron M., Camiade E., Raux G., Courtin P., Chapot-Chartier M. P., Pons J. L. (2005). Acd, a peptidoglycan hydrolase of *Clostridium difficile* with N-acetylglucosaminidase activity. *Microbiology* ***151:****2343-51.*

Dietrich C., Heuner K., Brand B.C., Hacker J., Steinert M. (2001). Flagellum of *Legionella pneumophila* positively affects the early phase of infection of eukaryotic host cells. *Infect Immun.* ***69****:2116-22.*

Dineen S.S., Villapakkam A.C., Nordman J.T., Sonenshein A.L. (2007). Repression of *Clostridium difficile* toxin gene expression by CodY. *Mol Microbiol.* ***66****:206-19.*

Dingle T.C., Mulvey G.L., Armstrong G.D. (2011). Mutagenic analysis of the *Clostridium difficile* flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters. *Infect Immun* ***79:****4061-7.*

Donskey C.J. (2010). Preventing transmission of *Clostridium difficile*: is the answer blowing in the wind? *Clin Infect Dis.* ***50****:1458-61.*

Doro F., Liberatori S., Rodriguez-Ortega M. J., Rinaudo C. D., Rosini R., Mora M., Scarselli M., Altindis E., D'Aurizio R., Stella M., Margarit I., Maione D., Telford J. L., Norais N., Grandi G. (2009). Surfome analysis as a fast track to vaccine discovery: identification of a novel protective antigen for Group B Streptococcus hypervirulent strain COH1. *Mol Cell Proteomics **8:**1728-37.*

Dorward D.W., Garon C.F. (1990). DNA is packaged within membrane-derived vesicles of Gram-negative but not Gram-positive bacteria. *Appl Environ Microbiol. **56**:1960-2.*

Drudy D., Harnedy N., Fanning S., Hannan M., Kyne L. (2007). Emergence and control of fluoroquinolone-resistant, toxin A-negative, toxin B-positive *Clostridium difficile*. *Infect Control Hosp Epidemiol. **28**:932-40.*

Dubberke E.R., Butler A.M., Yokoe D.S., Mayer J., Hota B., Mangino J.E., Khan Y.M., Popovich K.J., Fraser V.J. (2010). Multicenter study of *Clostridium difficile* infection rates from 2000 to 2006. *Infect Control Hosp Epidemiol. **31**:1030-7.*

.

Eaton K.A., Suerbaum S., Josenhans C., Krakowka S. (1996). Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. *Infect Immun. **64**:2445-8*.

Ellis T.N., Kuehn M.J. (2010). Virulence and immunomodulatory roles of bacterial outer membrane vesicles. *Microbiol Mol Biol Rev. **74**:81-94.*.

Emerson J.E., Reynolds C.B., Fagan R.P., Shaw H.A., Goulding D., Fairweather N.F. (2009). A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. *Mol Microbiol.**74**:541-56.*

Eveillard M., Fourel V., Barc M.C., Kernéis S., Coconnier M.H., Karjalainen T., Bourlioux P., Servin A.L. (1993). Identification and characterization of adhesive factors of *Clostridium difficile* involved in adhesion to human colonic enterocyte-like Caco-2 and mucus-secreting HT29 cells in culture. *Mol Microbiol. **7**:371-81.*

Fagan R.P., Albesa-Jové D., Qazi O., Svergun D.I., Brown K.A., Fairweather N.F. (2009). Structural insights into the molecular organization of the S-layer from *Clostridium difficile*. *Mol Microbiol. **71**:1308-22.*

Finkelstein R.A., Boesman-Finkelstein M., Holt P. (1983). *Vibrio cholerae* hemagglutinin/lectin/protease hydrolyzes fibronectin and ovomucin: F.M. Burnet revisited. *Proc Natl Acad Sci U S A **80:**1092-5.*

Freeman J., Bauer M.P., Baines S.D., Corver J., Fawley W.N., Goorhuis B., Kuijper E.J., Wilcox M.H. (2010). The changing epidemiology of *Clostridium difficile* infections. *Clin Microbiol Rev. **23**:529-49.*

Fritze D. (2004). Taxonomy of the genus Bacillus and related genera: the aerobic endospore-forming bacteria. *Phytopathology. **94**:1245-8.*

Gerding D.N. (2004). Clindamycin, cephalosporins, fluoroquinolones, and *Clostridium difficile*-associated diarrhea: this is an antimicrobial resistance problem. *Clin Infect Dis. **38**:646-8.*

Gerding D.N. (2009). *Clostridium difficile* 30 years on: what has, or has not, changed and why? *Int J Antimicrob Agents. 33 Suppl 1:S2-8.*

Geric B., Carman R.J., Rupnik M., Genheimer C.W., Sambol S.P., Lyerly D.M., Gerding D.N., Johnson S. (2006). Binary toxin-producing, large clostridial toxin-negative *Clostridium difficile* strains are enterotoxic but do not cause disease in hamsters. *J Infect Dis. **193**:1143-50.*

Grant C.C., Konkel M.E., Cieplak W. Jr, Tompkins L.S. (1993). Role of flagella in adherence, internalization, and translocation of *Campylobacter jejuni* in nonpolarized and polarized epithelial cell cultures. *Infect Immun. **61**:1764-71.*

Guillemet E., Cadot C., Tran S. L., Guinebretiere M. H., Lereclus D., Ramarao N. (2010). The InhA metalloproteases of *Bacillus cereus* contribute concomitantly to virulence. *J Bacteriol **192**:286-94.*

Gurung M., Moon D.C., Choi C.W., Lee J.H., Bae Y.C., Kim J., Lee Y.C., Seol S.Y., Cho D.T., Kim S.I., Lee J.C. (2011). *Staphylococcus aureus* produces membrane-derived vesicles that induce host cell death. *PLoS One **6**:e27958.*

He M., Miyajima F., Roberts P., Ellison L., Pickard D.J., Martin M.J., Connor T.R., Harris S.R., Fairley D., Bamford K.B., D'Arc S., Brazier J., Brown D., Coia J.E., Douce G., Gerding D., Kim H.J., Koh T.H., Kato H., Senoh M., Louie T., Michell S., Butt E., Peacock S.J., Brown N.M., Riley T., Songer G., Wilcox M., Pirmohamed M., Kuijper E., Hawkey P., Wren B.W., Dougan G., Parkhill J., Lawley T.D. (2012). Emergence and global spread of epidemic healthcare-associated *Clostridium difficile. Nat Genet. **45**:109-13.*

Henderson B., Nair S., Pallas J., Williams M.A. (2011). Fibronectin: a multidomain host adhesin targeted by bacterial fibronectin-binding proteins. *FEMS Microbiol Rev.**35**:147-200.*

Hennequin C., Collignon A., Karjalainen T. (2001a). Analysis of expression of GroEL (Hsp60) of *Clostridium difficile* in response to stress. *Microb Pathog.* ***31****:255-60.*

Hennequin C., Janoir C., Barc M.C., Collignon A., Karjalainen T. (2003). Identification and characterization of a fibronectin-binding protein from *Clostridium difficile. Microbiology.* ***149****:2779-87.*

Hennequin C., Porcheray F., Waligora-Dupriet A., Collignon A., Barc M., Bourlioux P., Karjalainen T. (2001b). GroEL (Hsp60) of *Clostridium difficile* is involved in cell adherence. *Microbiology*. ***147****:87-96.*

Hong S.W., Kim M.R., Lee E.Y., Kim J.H., Kim Y.S., Jeon S.G., Yang J.M., Lee B.J., Pyun B.Y., Gho Y.S., Kim Y.K. (2011). Extracellular vesicles derived from *Staphylococcus aureus* induce atopic dermatitis-like skin inflammation. *Allergy*. ***66****:351-9. .*

Jain S., Graham C., Graham R.L., McMullan G., Ternan N.G. (2011). Quantitative proteomic analysis of the heat stress response in *Clostridium difficile* strain 630. *J Proteome Res.****10****:3880-90*.

Janoir C., Pechine S., Grosdidier C., Collignon A. (2007). Cwp84, a surface-associated protein of *Clostridium difficile*, is a cysteine protease with degrading activity on extracellular matrix proteins. *J Bacteriol* ***189:****7174-80.*

Janvilisri T., Scaria J., Chang Y.F. (2010). Transcriptional profiling of *Clostridium difficile* and Caco-2 cells during infection. *J Infect Dis.* ***202****:282-90.*

Johnson S., Adelmann A., Clabots C.R., Peterson L.R., Gerding D.N. (1989). Recurrences of *Clostridium difficile* diarrhea not caused by the original infecting organism. *J Infect Dis.* ***159****:340-3.*

Kallen A.J., Thompson A., Ristaino P., Chapman L., Nicholson A., Sim B.T., Lessa F., Sharapov U., Fadden E., Boehler R., Gould C., Limbago B., Blythe D., McDonald L.C. (2009). Complete restriction of fluoroquinolone use to control an outbreak of *Clostridium difficil*e infection at a community hospital. *Infect Control Hosp Epidemiol.* ***30****:264-72.*

Karasawa T., Ikoma S., Yamakawa K., Nakamura S. **(1995).** A defined growth medium for *Clostridium difficile*. *Microbiology* ***141:****371-5.*

Karjalainen T., Waligora-Dupriet A.J., Cerquetti M., Spigaglia P., Maggioni A., Mauri P., Mastrantonio P. (2001) Molecular and genomic analysis of genes encoding surface-anchored proteins from *Clostridium difficile*. *Infect Immun.**69**:3442-6.*

Kesty N.C., Kuehn M.J. (2004). Incorporation of heterologous outer membrane and periplasmic proteins into *Escherichia coli* outer membrane vesicles. *J Biol Chem. **279**:2069-76.*

Kim J., Smathers S.A., Prasad P., Leckerman K.H., Coffin S., Zaoutis T. (2008). Epidemiological features of *Clostridium difficile*-associated disease among inpatients at children's hospitals in the United States, 2001-2006. *Pediatrics. **122**:1266-70.*

Kirby J.M., Ahern H., Roberts A.K., Kumar V., Freeman Z., Acharya K.R., Shone C.C. (2009). Cwp84, a surface-associated cysteine protease, plays a role in the maturation of the surface layer of Clostridium difficile. *J Biol Chem. **284**:34666-73.*

Klimpel K.R., Arora N., Leppla S.H. (1994). Anthrax toxin lethal factor contains a zinc metalloprotease consensus sequence which is required for lethal toxin activity. *Mol Microbiol **13:**1093-100.*

Klock H.E., Lesley S.A. (2009). The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. *Methods Mol Biol **498:**91-103.*

Kuehne S.A., Cartman S.T., Heap J.T., Kelly M.L., Cockayne A., Minton N.P. (2010). The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature **467:**711-3.*

Kuehne S.A., Heap J.T., Cooksley C.M., Cartman S.T., Minton N.P. (2011). ClosTron-mediated engineering of Clostridium. *Methods Mol Biol **765**:389-407.*

Kuipers E.J., Surawicz C.M. (2008). *Clostridium difficile* infection. *Lancet **371:**1486-8.*

Lamontagne F., Labbé A.C., Haeck O., Lesur O., Lalancette M., Patino C., Leblanc M., Laverdière M., Pépin J. (2007). Impact of emergency colectomy on survival of patients with fulminant *Clostridium difficile* colitis during an epidemic caused by a hypervirulent strain. *Ann Surg. **245**:267-72.*

Lawley T.D., Clare S., Walker A.W., Goulding D., Stabler R.A., Croucher N., Mastroeni P., Scott P., Raisen C., Mottram L., Fairweather N.F., Wren B.W., Parkhill J., Dougan G. (2009). Antibiotic treatment of *Clostridium difficile* carrier mice triggers a supershedder state, spore-mediated transmission, and severe disease in immunocompromised hosts. *Infect Immun* ***77:****3661-9.*

Lee E.Y., Choi D.Y., Kim D.K., Kim J.W., Park J.O., Kim S., Kim S.H., Desiderio D.M., Kim Y.K., Kim K.P., Gho Y.S. (2009). Gram-positive bacteria produce membrane vesicles: proteomics-based characterization of *Staphylococcus aureus*-derived membrane vesicles. *Proteomics.* ***9****:5425-36.*

Leonard J., Marshall J.K., Moayyedi P. (2007). Systematic review of the risk of enteric infection in patients taking acid suppression. *Am J Gastroenterol.* ***102****:2047-56.*

Lin Y.P., Kuo C.J., Koleci X., McDonough S.P., Chang Y.F. (2011). Manganese binds to *Clostridium difficile* Fbp68 and is essential for fibronectin binding. *J Biol Chem.* ***286****:3957-69.*

Linsky A., Gupta K., Lawler E.V., Fonda J.R., Hermos J.A. (2010). Proton pump inhibitors and risk for recurrent *Clostridium difficile* infection. Is *Clostridium difficile*-associated infection a potentially zoonotic and foodborne disease? *Arch Intern Med.* ***170****:772-8.*

Loo V.G., Poirier L., Miller M.A., Oughton M., Libman M.D., Michaud S., Bourgault A.M., Nguyen T., Frenette C., Kelly M., Vibien A., Brassard P., Fenn S., Dewar K., Hudson T.J., Horn R., René P., Monczak Y., Dascal A. (2005). A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. *N Engl J Med.* ***353****:2442-9.*

Lyras D., O'Connor J.R., Howarth P.M., Sambol S.P., Carter G.P., Phumoonna T., Poon R., Adams V., Vedantam G., Johnson S., Gerding D.N., Rood J.I. (2009). Toxin B is essential for virulence of *Clostridium difficile*. *Nature* ***458:****1176-9.*

Maeda H. (1996). Role of microbial proteases in pathogenesis. *Microbiol Immunol* ***40:****685-99.*

Mah T.F., O'Toole G.A. (2001). Mechanisms of biofilm resistance to antimicrobial agents. *Trends Microbiol* ***9:****34-9.*

Makarova K.S., Aravind L., Koonin E.V. (1999). A superfamily of archaeal, bacterial, and eukaryotic proteins homologous to animal transglutaminases. *Protein Sci.* ***8****:1714-9.*

Matsumoto K. (2004). Role of bacterial proteases in pseudomonal and serratial keratitis. *Biol Chem.* ***385****:1007-16.*

McDonald L.C., Killgore G.E., Thompson A., Owens R.C. Jr, Kazakova S.V., Sambol S.P., Johnson S., Gerding D.N. (2005). An epidemic, toxin gene-variant strain of *Clostridium difficile*. *N Engl J Med.* ***353****:2433-41.*

McSweegan E., Walker R.I. (1986). Identification and characterization of two *Campylobacter jejuni* adhesins for cellular and mucous substrates. *Infect Immun.****53****:141-8.*

Miura M., Kato H., Matsushita O. (2011). Identification of a novel virulence factor in *Clostridium difficile* that modulates toxin sensitivity of cultured epithelial cells. *Infect Immun* ***79:****3810-20.*

Molla A., Tanase S., Hong Y. M., Maeda H. (1988). Interdomain cleavage of plasma fibronectin by zinc-metalloproteinase from *Serratia marcescens*. *Biochim Biophys Acta* ***955:****77-85.*

Monot M., Boursaux-Eude C., Thibonnier M., Vallenet D., Moszer I., Medigue C., Martin-Verstraete I., Dupuy B. (2011). Reannotation of the genome sequence of *Clostridium difficile* strain 630. *J Med Microbiol* ***60:****1193-9.*

Mukherjee K., Karlsson S., Burman L.G., Akerlund T. (2002). Proteins released during high toxin production in *Clostridium difficile*. *Microbiology* ***148:****2245-53.*

Navarre W.W., Schneewind O. (1999). Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev.* ***63****:174-229.*

Nelson D.C., Garbe J., Collin M. (2011). Cysteine proteinase SpeB from *Streptococcus pyogenes* - a potent modifier of immunologically important host and bacterial proteins. *Biol Chem* ***392:****1077-88.*

Niesen F.H., Berglund H., Vedadi M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* ***2:****2212-21.*

Nobbs A.H., Lamont R. J., Jenkinson H.F. (2009). Streptococcus adherence and colonization. *Microbiol Mol Biol Rev **73:**407-50*

O'Connor J.R., Johnson S., Gerding D.N. (2009). *Clostridium difficile* infection caused by the epidemic BI/NAP1/027 strain. *Gastroenterology **136:**1913-24.*

Olling A., Seehase S., Minton N.P., Tatge H., Schröter S., Kohlscheen S., Pich A., Just I., Gerhard R. (2012). Release of TcdA and TcdB from *Clostridium difficile* cdi 630 is not affected by functional inactivation of the tcdE gene. *Microb Pathog. **52**:92-100.*

O'Neill G.L., Beaman M.H., Riley T.V. (1991). Relapse versus reinfection with *Clostridium difficile. Epidemiol Infect. **107**:627-35.*

Pacheco L.G., Slade S.E., Seyffert N., Santos A.R., Castro T.L., Silva W.M., Santos A.V., Santos S.G., Farias L.M., Carvalho M.A., Pimenta A.M., Meyer R., Silva A., Scrivens J.H., Oliveira S.C., Miyoshi A., Dowson C.G., Azevedo V. (2011). A combined approach for comparative exoproteome analysis of *Corynebacterium pseudotuberculosis. BMC Microbiol **11:**12.*

Péchiné S., Denève C., Le Monnier A., Hoys S., Janoir C., Collignon A. (2011). Immunization of hamsters against *Clostridium difficile* infection using the Cwp84 protease as an antigen. *FEMS Immunol Med Microbiol. **63**:73-81.*

Pechine S., Gleizes A., Janoir C., Gorges-Kergot R., Barc M.C., Delmee M., Collignon A. (2005a). Immunological properties of surface proteins of *Clostridium difficile*. *J Med Microbiol **54:**193-6.*

Péchiné S., Janoir C., Collignon A. (2005b). Variability of *Clostridium difficile* surface proteins and specific serum antibody response in patients with *Clostridium difficile*-associated disease. *J Clin Microbiol. **43**:5018-25.*

Pépin J., Valiquette L., Alary M.E., Villemure P., Pelletier A., Forget K., Pépin K., Chouinard D. (2004). *Clostridium difficile*-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. *CMAJ. **171**:466-72.*

Poilane I., Karjalainen T., Barc M.C., Bourlioux P., Collignon A. (1998). Protease activity of *Clostridium difficile* strains. *Can J Microbiol **44:**157-61.*

Poxton I.R., McCoubrey J., Blair G. (2001). The pathogenicity of *Clostridium difficile*. *Clin Microbiol Infect. **7**:421-7.*

Ravipaty S., Reilly J.P. (2010). Comprehensive characterization of methicillin-resistant *Staphylococcus aureus* subsp. aureus COL secretome by two-dimensional liquid chromatography and mass spectrometry. *Mol Cell Proteomics **9:**1898-919.*

Redelings M.D., Sorvillo F., Mascola L.(2007). Increase in *Clostridium difficile*-related mortality rates, United States, 1999-2004. *Emerg Infect Dis.**13**:1417-9.*

Reynolds C.B., Emerson J.E., de la Riva L., Fagan R.P., Fairweather N.F. (2011). The *Clostridium difficile* cell wall protein CwpV is antigenically variable between strains, but exhibits conserved aggregation-promoting function. *PLoS Pathog **7:**e1002024.*

Ricciardi R., Rothenberger D.A., Madoff R.D., Baxter N.N. (2007). Increasing prevalence and severity of *Clostridium difficile* colitis in hospitalized patients in the United States. *Arch Surg. **142**:624-31.*

Rigden D.J., Galperin M.Y. (2008). Sequence analysis of GerM and SpoVS, uncharacterized bacterial 'sporulation' proteins with widespread phylogenetic distribution. *Bioinformatics **24:**1793-7.*

Rivera J., Cordero R.J., Nakouzi A.S., Frases S., Nicola A., Casadevall A. (2010). *Bacillus anthracis* produces membrane-derived vesicles containing biologically active toxins. *Proc Natl Acad Sci U S A. **107**:19002-7.*

Rodriguez-Ortega M. J., Norais N., Bensi G., Liberatori S., Capo S., Mora M., Scarselli M., Doro F., Ferrari G., Garaguso I., Maggi T., Neumann A., Covre A., Telford J. L., Grandi G. (2006). Characterization and identification of vaccine candidate proteins through analysis of the group A Streptococcus surface proteome. *Nat Biotechnol **24:**191-7.*

Rodriguez-Palacios A., Stämpfli H.R., Duffield T., Peregrine A.S., Trotz-Williams L.A., Arroyo L.G., Brazier J.S., Weese J.S. (2006). *Clostridium difficile* PCR ribotypes in calves, Canada. *Emerg Infect Dis. **12**:1730-6.*

Rupnik M. (2007). Is *Clostridium difficile*-associated infection a potentially zoonotic and foodborne disease? *Clin Microbiol Infect.**13**:457-9.*

Rupnik M., Wilcox M.H., Gerding D.N. (2009). *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. *Nat Rev Microbiol **7:**526-36.*

Sammons R.L., Slynn G.M., Smith D.A. (1987). Genetical and molecular studies on gerM, a new developmental locus of *Bacillus subtilis*. *J Gen Microbiol **133:**3299-312*.

Sandolo C., Péchiné S., Le Monnier A., Hoys S., Janoir C., Coviello T., Alhaique F., Collignon A., Fattal E., Tsapis N. (2011). Encapsulation of Cwp84 into pectin beads for oral vaccination against *Clostridium difficile*. *Eur J Pharm Biopharm. **79**:566-73.*

Sára M., Sleytr U.B. (2000). S-Layer proteins. *J Bacteriol. **182**:859-68.*

Sarker M.R., Paredes-Sabja D. (2012). Molecular basis of early stages of *Clostridium difficile* infection: germination and colonization. *Future Microbiol. **7**:933-43.*

Savariau-Lacomme M.P., Lebarbier C., Karjalainen T., Collignon A., Janoir C. (2003). Transcription and analysis of polymorphism in a cluster of genes encoding surface-associated proteins of *Clostridium difficile*. *J Bacteriol. **185**:4461-70.*

Schirmer J., Aktories K. (2004). Large clostridial cytotoxins: cellular biology of Rho/Ras-glucosylating toxins. *Biochim Biophys Acta. **1673**:66-74.*

Schwan C., Stecher B., Tzivelekidis T., van Ham M., Rohde M., Hardt W.D., Wehland J., Aktories K. (2009). *Clostridium difficile* toxin CDT induces formation of microtubule-based protrusions and increases adherence of bacteria. *PLoS Pathog **5:**e1000626.*

Sebaihia M., Wren B.W., Mullany P., Fairweather N.F., Minton N., Stabler R., Thomson N.R., Roberts A.P., Cerdeno-Tarraga A.M., Wang H., Holden M.T., Wright A., Churcher C., Quail M.A., Baker S., Bason N., Brooks K., Chillingworth T., Cronin A., Davis P., Dowd L., Fraser A., Feltwell T., Hance Z., Holroyd S., Jagels K., Moule S., Mungall K., Price C., Rabbinowitsch E., Sharp S., Simmonds M., Stevens K., Unwin L., Whithead S., Dupuy B., Dougan G., Barrell B., Parkhill J. (2006). The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nat Genet **38**:779-86.*

Seddon S.V., Borriello S.P. (1992). Proteolytic activity of *Clostridium difficile*. *J Med Microbiol **36:**307-11.*

Seddon S.V., Hemingway I., Borriello S.P. (1990). Hydrolytic enzyme production by *Clostridium difficile* and its relationship to toxin production and virulence in the hamster model. *J Med Microbiol* ***31:****169-74.*

Sengupta N., Alam S.I., Kumar B., Kumar R.B., Gautam V., Kumar S., Singh L. (2010). Comparative proteomic analysis of extracellular proteins of *Clostridium perfringens* type A and type C strains. *Infect Immun* ***78:****3957-68.*

Setlow P. (2007). I will survive: DNA protection in bacterial spores. *Trends Microbiol.* ***15****:172-80.*

Slynn G.M., Sammons R.L., Smith D.A., Moir A., Corfe B.M. (1994). Molecular genetical and phenotypical analysis of the gerM spore germination gene of *Bacillus subtilis* 168. *FEMS Microbiol Lett* ***121:****315-20.*

Songer J.G., Anderson M.A. (2006). *Clostridium difficile*: an important pathogen of food animals. *Anaerobe*. ***12****:1-4.*

Sundriyal A., Roberts A.K., Ling R., McGlashan J., Shone C.C., Acharya K.R. (2010). Expression, purification and cell cytotoxicity of actin-modifying binary toxin from *Clostridium difficile*. *Protein Expr Purif.* ***74****:42-8.*

Takeoka A., Takumi K., Koga T., Kawata T. (1991). Purification and characterization of S layer proteins from *Clostridium difficile* GAI 0714. *J Gen Microbiol.* ***137****:261-7.*

Tam Dang T.H., Fagan R.P., Fairweather N.F., Tate E.W. (2012). Novel inhibitors of surface layer processing in *Clostridium difficile*. *Bioorg Med Chem.* ***20****:614-21*.

Tasteyre A., Barc M.C., Collignon A., Boureau H., Karjalainen T. (2001a). Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. *Infect Immun* ***69:****7937-40.*

Tasteyre A., Karjalainen T., Avesani V., Delmée M., Collignon A., Bourlioux P., Barc M.C. (2001b). Molecular characterization of *fliD* gene encoding flagellar cap and its expression among *Clostridium difficile* isolates from different serogroups. *J Clin Microbiol.* ***39****:1178-83.*

Teo J.W., Zhang L.H., Poh C.L. (2003). Cloning and characterization of a metalloprotease from *Vibrio harveyi* strain AP6. *Gene* ***303:****147-56.*

Timpl R., Brown J.C. (1996). Supramolecular assembly of basement membranes. *Bioessays. **18**:123-32.*

Tonello F., Montecucco C. (2009). The anthrax lethal factor and its MAPK kinase-specific metalloprotease activity. *Mol Aspects Med **30:**431-8.*

van der Flier M., Chhun N., Wizemann T.M., Min J., McCarthy J.B., Tuomanen E.I. (1995). Adherence of *Streptococcus pneumoniae* to immobilized fibronectin. *Infect Immun.**63**:4317-22.*

Vedantam G., Clark A., Chu M., McQuade R., Mallozzi M., Viswanathan V.K. (2012). *Clostridium difficile* infection: toxins and non-toxin virulence factors, and their contributions to disease establishment and host response. *Gut Microbes. **3**:121-34.*

Voth D.E., Ballard J.D. (2005). *Clostridium difficile* toxins: mechanism of action and role in disease. *Clin Microbiol Rev **18:**247-63.*

Wai S.N., Lindmark B., Söderblom T., Takade A., Westermark M., Oscarsson J., Jass J., Richter-Dahlfors A., Mizunoe Y., Uhlin B.E. (2003). Vesicle-mediated export and assembly of pore-forming oligomers of the enterobacterial ClyA cytotoxin. *Cell. **115**:25-35.*

Waligora A.J., Hennequin C., Mullany P., Bourlioux P., Collignon A., Karjalainen T. (2001). Characterization of a cell surface protein of *Clostridium difficile* with adhesive properties. *Infect Immun **69:**2144-53.*

Waligora A.J., Barc M.C., Bourlioux P., Collignon A., Karjalainen T. (1999). *Clostridium difficile* cell attachment is modified by environmental factors. *Appl Environ Microbiol. **65**:4234-8.*

Warny M., Pepin J., Fang A., Killgore G., Thompson A., Brazier J., Frost E., McDonald L.C. (2005). Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. *Lancet **366:**1079-84.*

Wright A., Drudy D., Kyne L., Brown K., Fairweather N.F. (2008). Immunoreactive cell wall proteins of *Clostridium difficile* identified by human sera. *J Med Microbiol **57:**750-6.*

Wright A., Wait R., Begum S., Crossett B., Nagy J., Brown K., Fairweather N. (2005). Proteomic analysis of cell surface proteins from *Clostridium difficile*. *Proteomics **5:**2443-52.*

Wüst J., Sullivan N.M., Hardegger U., Wilkins TD. (1982). Investigation of an outbreak of antibiotic-associated colitis by various typing methods. *J Clin Microbiol.* ***16****:1096-101*.